TESTS FOR HARMONIC COMPONENTS

IN BIVARIATE TIME SERIES

by

Donald F. Morrison

University of North Carolina

Sponsored by the Office of Naval Research under the contract for research in probability and statistics at Chapel Hill. Reproduction in whole or in part is permitted for any purpose of the United States government.

Institute of Statistics
Mimeograph Series No. 165
April, 1957

ACKNOWLEDGEMENT

I wish to express my deep appreciation to Professor Harold Hotelling for introducing me to the many unsolved problems of periodogram analysis, and in particular, for his suggestion of the need for a bivariate test of significant periods in parallel time series.

I am especially indebted to Mrs. Hilda Kattsoff for typing these pages under the difficult conditions imposed by my absence from Chapel Hill.

Finally, I wish to convey my thanks to the Office of Naval Research for their generous support of my work while at the University of North Carolina.

TABLE OF CONTENTS

INTRODUCTION[1]

The purpose of this thesis is that of setting forth a test for sinusoidal components of equal period in two parallel time series. A test based upon a "generalized intensity" in the form of a determinant is proposed and the exact distribution of the greatest of these intensities is easily obtained from existing distribution theory.

Chapter I is a review of the least squares analysis of periodicities in a single time series. Schuster's [13][2] and Walker's [17] tests are discussed, and Fisher's 1929 [5] distribution of the ratio of the greatest intensity to the sum of the intensities is given in detail. Some studies of the effect of the model period being situated between two adjacent trial periods are presented. The power function for Walker's test is obtained, and is shown to be unbiassed.

A bivariate test for simple sinusoids in each of two concurrent series is presented in Chapter II. A general theorem on the estimation of parameters in multivariate regression is given, and the estimates of the coefficients for the case of unequal model periods are obtained to show their dependence upon certain nuisance parameters. An extremely simple exact test based on the "intensity determinant" follows, with a short discussion of some of its pro-

---

1. Sponsored by the Office of Naval Research under the contract for research in probability and statistics at Chapel Hill. Reproduction in whole or in part is permitted for any purpose of the United States government.

2. The numbers in square brackets refer to the bibliography listed at the end.

perties. Other possible test statistics are mentioned, none of which appears to give rise to such a simple distribution as the intensity determinant.

The expectation of the intensity determinant at a specified point is obtained under a variety of conditions and model series in Chapter III. The expectations of the twelve possible combinations of equal, unequal, and non-existent periods in the two series with equal or unequal trial periods and two types of estimates of the parameters based on equal or unequal model periods, are computed at a point and compared.

CHAPTER I

TESTS OF SIGNIFICANCE FOR PERIODICITIES

1.1. Introduction

This chapter is an expository review of tests of significance for a simple sinusoidal component in a single series of observations ordered in time. No attempt is made to summarize the entire field of harmonic analysis as it was set forth mainly by Fourier in 1822. Rather, this chapter will be restricted to the concepts arising from Sir Arthur Schuster's [13] definition of a periodogram, and its ramifications by later workers in modern statistical theory.

1.2. The Least Squares Model

We hypothesize that the observations arise from the structure

(1.2.1) $$u_t = A_o + D\sin t(\theta - \theta_o) + \varepsilon_t, \quad t = 1, \ldots, N.$$

For estimation purposes, ε_t is a random variable satisfying the usual least squares conditions

(1.2.2) $$\begin{array}{l} E(\varepsilon_t) = 0 \\ E(\varepsilon_r \varepsilon_s) = \delta_{rs}\sigma^2 \end{array}, \text{ where } \delta_{rs} = \begin{array}{l} 1, \quad r = s \\ 0, \quad r \neq s \end{array} .$$

To make tests of hypotheses concerning the parameters of (1.2.1), it is further required that ε_t is a normally distributed random variable. This condition will be denoted by

(1.2.3) $$\varepsilon_t : N(0, \sigma^2), \; t = 1, \ldots, N .$$

The phase parameter Θ_o entering non-linearly into the model may be changed into a more tractable form by writing the model equations as

(1.2.4) $$u_t = C + D\sin t\Theta \cos t\Theta_o - D\cos t\Theta \sin t\Theta_o + \varepsilon_t \quad ,$$

or as,

(1.2.5) $$u_t = C + A\cos t\Theta + B \sin t\Theta + \varepsilon_t \quad ;$$

Θ_o may be determined by the relation $\Theta_o = \tan^{-1}(B/A)$. While it is sometimes desirable to put confidence bounds on the phase, we shall be concerned only with the problem of estimating Θ. In all further work, (1.2.5) will be taken as the univariate model for a simple sinusoidal component.

1.3. The Least Squares Estimates of the Parameters C, A, B, and Θ

Unless otherwise specified, in all further work the restrictions will be made that N is odd, and Θ is of the form

(1.3.1) $$\Theta = \frac{2\pi k}{N} \quad , k = 1, \ldots, \frac{N-1}{2} = n \quad .$$

This restriction on Θ is made to use the exact orthogonality of certain trigonometric functions when summed over suitable intervals. A variety of such sums needed in this study is given in the Appendix.

For a time series of N observations, we wish to estimate the parameters C, A, B, and Θ so that the residual sum of squares

(1.3.2) $$\sum_{t=1}^{N} (u_t - A\cos t\Theta - B\sin t\Theta - C)^2$$

is a minimum. Making the matrix definitions

$$(1.3.3)\quad x = \begin{bmatrix} 1 & \cos\theta & \sin\theta \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & \cos N\theta & \sin N\theta \end{bmatrix}, \quad \underline{y} = \begin{bmatrix} u_1 \\ \cdot \\ \cdot \\ \cdot \\ u_N \end{bmatrix}, \quad b = \begin{bmatrix} \hat{C} \\ \hat{A} \\ \hat{B} \end{bmatrix},$$

the least squares estimates of C, A, and B are given by

$$b = (x'x)^{-1}(x'y)$$

$$(1.3.4)\quad = \begin{bmatrix} N & 0 & 0 \\ 0 & \frac{N}{2} & 0 \\ 0 & 0 & \frac{N}{2} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{N} u_t \\ \sum_{t=1}^{N} u_t \cos t\theta \\ \sum_{t=1}^{N} u_t \sin t\theta \end{bmatrix},$$

where the diagonal matrix $(x'x)$ has resulted from the use of the Appendix sums (A.2.1) - (A.2.5).

Hence, the least square estimates are simply

$$(1.3.5)\quad \begin{aligned} \hat{C} &= \frac{1}{N} \sum_{t=1}^{N} u_t \\ \hat{A} &= \frac{2}{N} \sum_{t=1}^{N} u_t \cos t\theta \\ \hat{B} &= \frac{2}{N} \sum_{t=1}^{N} u_t \sin t\theta \end{aligned} .$$

To find the least square estimate of θ, consider the error sum of squares (1.3.2). Using the three estimates just obtained, this may be written as

$$\text{(1.3.6)} \qquad \begin{aligned} SSE &= \sum_{t=1}^{N} (u_t - \hat{C} - \hat{A}\cos t\theta - \hat{B}\sin t\theta)^2 \\ &= \sum_{t=1}^{N} u_t^2 - \left(\frac{\sum_{t=1}^{N} u_t}{N}\right)^2 - \frac{N}{2}(\hat{A}^2 + \hat{B}^2) \end{aligned} \quad .$$

This is clearly at a minimum when the quantity

$$\text{(1.3.7)} \qquad I = \hat{A}^2 + \hat{B}^2$$

is at a maximum. I is called the "intensity" associated with a particular value of θ. Since we have restricted our attention to θ's of the form $\frac{2\pi k}{N}$, $k=1, 2, \ldots, n$, our estimate of the true period is N/k, where k is the integer from the set $1, \ldots, n$ that makes I a maximum. We have thus connected Schuster's original periodogram concept with the least squares criterion.

1.4. Tests of Significance of Trial Periods

Schuster [13] is apparently the first to have considered a test for the significance of a single trial period. He proposed that the n intensities be computed, and then be plotted on a chart against $k = 1, \ldots, n$. The period is taken to be N/k_o, where k_o is the argument corresponding to the greatest periodogram ordinate. Although not set forth in this more modern notation, Schuster went further to propose a test of the hypothesis

$$\text{(1.4.1)} \qquad H_o\colon\ A^2 + B^2 = 0 \quad ,$$

against the alternative

$$H_a: A^2 + B^2 = k \neq 0 \quad . \tag{1.4.2}$$

Since A and B are real numbers, this test is equivalent to testing whether A and B are both 0, and hence whether a periodic component exists in the model.

We shall obtain the distribution of the sample intensity $\hat{A}^2 + \hat{B}^2$ under the null hypothesis (1.4.1) in modern terminology, then discuss in what respect this differs from Schuster's original result. Since the ε_t are distributed with zero mean and variance σ^2, it follows from the properties of the trigonometric sums (A.2.4)-(A.2.6) of the Appendix that

$$\hat{A} = \frac{2}{N} \sum_{t=1}^{N} \cos\ t\theta \tag{1.4.3}$$

and

$$\hat{B} = \frac{2}{N} \sum_{t=1}^{N} \sin t\theta$$

each have a normal distribution with zero mean and variance $\frac{2\sigma^2}{N}$, and from the fact that $\sum_{t=1}^{N} \sin t\theta \cos t\theta = 0$, they are distributed independently of each other. Thus,

$$x = \frac{N(\hat{A}^2 + \hat{B}^2)}{2\sigma^2} \tag{1.4.4}$$

has a χ^2 distribution with two degrees of freedom,

$$f(x)dx = \frac{1}{2} e^{-\frac{x}{2}} dx, \quad x \geq 0 \quad . \tag{1.4.5}$$

The probability that $x = \frac{N(\hat{A}^2+\hat{B}^2)}{2\sigma^2}$ will exceed the value x' by chance under the null hypothesis of a strictly random normal series is given by the cumulative distribution

$$(1.4.6) \qquad P(x > x') = e^{-\frac{x'}{2}} \quad .$$

This expression is generally referred to as "Schuster's Probability" by contemporary writers on time series.

It will be interesting to compare this expression with Schuster's actual 1898 [13] paper. Schuster took as a measure of intensity the square root of our expression (1.3.7), denoting it by

$$(1.4.7) \qquad r_1 = \sqrt{a_1^2 + b_1^2} \quad .$$

He then divided this by $a_o = \frac{1}{p} \sum_{t=1}^{N} u_t$, for p = length of trial period, to give his measure $\rho = \frac{r_1}{a_o}$. Disregarding the fact that a_o is a random variable, Schuster states that the probability element of ρ is

$$(1.4.8) \qquad H(\rho)d\rho = \frac{n}{2}\rho \exp(-\frac{n\rho^2}{4})d\rho \quad .$$

From this he shows that $E(\rho) = \sqrt{\frac{\pi}{N}}$, $E(\rho^2) = \frac{4}{n}$, and $P(\rho > \rho') = \exp(-\frac{n\rho'^2}{4})$. While these are valid for a_o constant, it must not have been apparent to Schuster that he was dealing with the ratio of two random variables, and not a single one.

Sir Gilbert Walker [17] noted in 1914 that Schuster's test failed to allow for the selection of the greatest intensity from a set of n, and gave the probability that the greatest intensity x should exceed x' as

$$(1.4.9) \qquad P(x > x') = 1 - \left[1 - e^{-\frac{x'}{2}}\right]^n ,$$

where $x = \dfrac{N(\hat{A}^2+\hat{B}^2)}{2\sigma^2}$ provided all the n statistics x are independent. The independence condition is satisfied if all these n trial periods are submultiples of the number N of observations. Tables of this probability are available in David, [3]. We shall discuss the power function of Walker's test in Section 1.8.

1.5. <u>Fisher's Distribution of the Ratio of the Greatest of a Set of Mean Squares to their Sum</u>

It remained for Sir Ronald Fisher to obtain the distribution for the general case of σ^2 unknown. In his 1929 paper [5], Fisher found the exact distribution of the ratio

$$(1.5.1) \qquad g = \frac{\max(\hat{A}_k^2+\hat{B}_k^2)}{(\hat{A}_1^2+\hat{B}_1^2) + \dots + (\hat{A}_n^2+\hat{B}_n^2)}$$

through the ingenious use of geometrical methods. We shall present here a derivation in the same manner communicated to the author by Professor Harold Hotelling, and in much greater detail than the original

1929 paper. An analytical derivation of the distribution of the r-th greatest intensity ratio has been given by Whittle [18], p. 99, through the use of the Lévy inversion theorem for the characteristic function of the joint distribution of the intensities. Cochran [?] also obtained the distribution more straightforwardly by transforming the n different g's to independent variates, then operating upon their joint density.

The cumulative form of Fisher's distribution for the ratio g given in (1.5.1) is

$$P_F(g > g') = \sum_{j=1}^{m} \binom{n}{j} (-)^{j+1} (1 - jg)^{n-1} \quad , \tag{1.5.1}$$

where $m = [g^{-1}]$, i.e., the largest integer $\leq g^{-1}$. This condition is required in order that the non-zero terms $(1 - jg)^{n-1}$ are always positive.

We shall now obtain Fisher's distribution geometrically. The joint density of the n quantities

$$x_i = \frac{N}{2\sigma^2} (\hat{A}_i^2 + \hat{B}_i^2), \ i = 1, 2, \ldots, n \quad ,$$

is

$$f(x_1, \ldots, x_n) = 2^{-n} \exp\left\{ - \sum_{i=1}^{n} x_i \right\} \quad . \tag{1.5.3}$$

It is apparent that the density is constant over hyperplanes of dimension $(n-1)$, and g may vary from n^{-1}, for all intensities equal, to 1, for all but one equal to 0. As g varies over this range, the portions of the hyperplane in which a given x_i is greatest may overlap, so that the distribution has discontinuities for certain definite values of g. The boundaries of these portions intersect in the following sets of points:

$$
(1.5.4) \qquad
\begin{array}{l}
x_1 = x_2 = x_3 \quad \cdot \quad = x_n = n^{-1} = g \\
x_2 = 0;\ x_1 = x_3 \ \cdot = x_n = (n-1)^{-1} = g \\
x_2 = x_3 = 0;\ x_1 = \ \ldots = x_n = (n-2)^{-1} = g \\
\cdot \quad \cdot \quad \cdot \qquad\qquad \cdot \quad \cdot \quad \cdot \\
\cdot \quad \cdot \quad \cdot \qquad\qquad \cdot \quad \cdot \quad \cdot \\
\cdot \quad \cdot \quad \cdot \qquad\qquad \cdot \quad \cdot \quad \cdot \\
x_2 = x_3 = x_4 \quad \ldots x_n = 0;\ x_1 = 1 = g
\end{array}
\quad .
$$

Using this knowledge of the discontinuities at these vertices, we may write the cumulative distribution in the form

$$
(1.5.5) \qquad P_F = \sum_{x=1}^{k} \alpha_x (1 - xg)^{n-1} \quad ,
$$

where $k = [g^{-1}]$, i.e., the greatest integer $\leq g^{-1}$. Since $g \geq 1^{-1}$, $P\ (n^{-1}) = 1$. Thus,

$$
(1.5.6) \qquad 1 = \sum_{x=1}^{n-1} \alpha_x (1 - \frac{x}{n})^{n-1} \quad .
$$

Similarly, the first (n-1) derivatives of the cumulative distribution function must vanish at $g = n^{-1}$. This requires that

$$
(1.5.7) \qquad 0 = \sum_{x=1}^{n-1} \alpha_x x^r (1 - \frac{x}{n})^{n-r-1}, \ r = 1, \ldots, n-2 \quad .
$$

We are thus provided with a system of $(n-1)$ equations in the $(n-1)$ unknowns $\alpha_1, \ldots, \alpha_{n-1}$. The matrix of this system is

$$(1.5.8) \qquad \begin{bmatrix} (1-\frac{1}{n})^{n-1} & (1-\frac{2}{n})^{n-1} & \cdots & (1-\frac{n-1}{n})^{n-1} \\ (1-\frac{1}{n})^{n-2} & 2(1-\frac{2}{n})^{n-2} & & (n-1)(1-\frac{n-1}{n})^{n-2} \\ \vdots & \vdots & & \vdots \\ (1-\frac{1}{n}) & 2^{n-2}(1-\frac{2}{n}) & & (n-1)^{n-2}(1-\frac{n-1}{n}) \end{bmatrix}$$

We wish to show that the rank of this matrix is $n-1$, hence that there exists a unique solution $(\alpha_1, \ldots, \alpha_{n-1})$ to the above equations. After some minor changes, the matrix becomes

$$(1.5.9) \qquad \begin{bmatrix} (\frac{n-1}{n})^{n-1} & \cdots & \frac{1}{n^{n-1}} \\ (\frac{n-1}{n})^{n-2} & & \\ \vdots & \ddots & \vdots \\ (\frac{n-1}{n}) & & (n-1)^{n-2} \end{bmatrix}$$

We may factor n^{-n+i} from the ith row for all i, and then multiply the jth column by $(n-j)^{1-n}$ for all j. This leaves

$$(1.5.10)\qquad \text{const.}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \frac{1}{n-1} & \frac{2}{n-2} & \cdots & (n-1) \\ \vdots & & & \vdots \\ \left(\frac{1}{n-1}\right)^{n-2} & \left(\frac{2}{n-2}\right)^{n-2} & \cdots & (n-1)^{n-2} \end{bmatrix} .$$

But this is a Vandermonde matrix with rank $n-1$, since none of the basic elements $i(n-i)^{-1}$ are equal. Hence, the system has a unique solution.

We shall state the solution as

$$(1.5.11)\qquad \alpha_x = (-)^{x-1}\binom{n}{x} ,$$

and then show by finite difference operations that this must be the unique one. An attempt was made to find an inverse to (1.5.8), but because of the asymmetry, no simply patterened inverse appears to be readily available. From the equations (1.5.6) and (1.5.7) ,

$$(1.5.12)\qquad \sum_{x=1}^{n-1}(-)^{x-1}\binom{n}{x}x^r\left(1-\frac{x}{n}\right)^{n-r-1} = \delta_{or},\ r = 0,\ 1,\ \ldots\ n-2 ,$$

where δ_{or} is the Kronecker δ. Let

$$(1.5.13)\qquad u_x = x^r\left(1-\frac{x}{n}\right)^{n-r-1} .$$

Then, the n-th difference, $\Delta^n u_x$, of this (n-1)-degree polynomial is zero. We further employ the "shift operator" E such that

$$E\ f(x) = f(x+1) = (1+\Delta)\ f(x) \quad . \tag{1.5.14}$$

Thus, $\Delta^n u_x = (-)^n(1-E)^n u_x$. Expanding the binomial and setting $x=0$, we have

$$0 = (-)^n(u_o - nu_1 + \binom{n}{2}u_2 - \dots + (-)^n u_n) \quad . \tag{1.5.15}$$

Since $u_n = 0$ by (1.5.13), the last term in this expansion can be omitted. Cancelling the $(-)^n$ as well, and rearranging terms,

$$u_o = nu_1 - \binom{n}{2}u_2 + \dots - (-)^{n-1}nu_{n-1}$$

$$= \sum_{x=1}^{n-1} (-)^{x-1}\binom{n}{x}u_x$$

$$= \sum_{x=1}^{n-1} (-1)^{x-1}\binom{n}{x}x^r(1-\frac{x}{n})^{n-r-1} \quad .$$

For $r=0$, $u_x = (1-\frac{x}{n})^{n-1}$, so that $u_o = 1$. For $r > 0$, $u_o \equiv 0$. Hence

$(-)^{x-1}\binom{n}{x} = \alpha_x$ is the unique solution of the system of equations.

Fisher's cumulative distribution becomes

$$P(g' > g) = \sum_{g=1}^{[\frac{1}{g}]} \binom{n}{j}(-)^{j+1}(1-jg)^{n-1} \quad , \tag{1.5.17}$$

as was to be shown.

Fisher [7] considered the equivalence of this distribution with the probability obtained by W. L. Stevens [15] for a certain geometrical problem. Fisher also noted that Stevens' more general result could be used to obtain the distribution of the second greatest ratio of intensities in the set of n. Fisher has made reference to this distribution in his 1939 paper [6] on the distributions of statistics arising from non-linear equations, and considered therein the mean of the distribution for large n. Since these results will not be required in the later chapters, we shall not consider them here.

1.6. An Identity Involving the Sum of the Intensities

We shall now establish in two ways a special case of Bessel's Inequality:

$$(1.6.1) \qquad \sum_{k=1}^{n} (\hat{A}_k^2+\hat{B}_k^2) = \frac{2}{N} \sum_{t=1}^{N} (u_t - \bar{u})^2 \text{ , } N = 2n + 1 \quad .$$

This enables us to find the sum of all n intensities without the laborious computation associated with each.

This identity can be shown most elegantly as follows: The residual sum of squares for the complete Fourier sequence is identically zero, or,

$$(1.6.2) \quad \sum_{t=1}^{N} (u_t-\hat{C}-\hat{A}_1 \cos \frac{2\pi t}{N} - \hat{B}_1 \sin \frac{2\pi t}{N} - \ldots -\hat{A}_n \cos\frac{2\pi n}{N}t-\hat{B}_n \sin\frac{2\pi nt}{N})^2$$

$$\equiv 0 \quad .$$

If we expand the summation and note from the Appendix that

$\sum_{t=1}^{N} \sin^2 t\theta = \frac{N}{2}$, $\sum_{t=1}^{N} \cos^2 t\theta = \frac{N}{2}$, $\sum_{t=1}^{N} \sin t\theta \cos t\theta = 0$, this becomes

$$\sum_{t=1}^{N} (u_t-\overline{u})^2-(\hat{A}_1^2+\hat{B}_n^2)\,\frac{N}{2} \equiv 0$$

Thus,

$$\frac{2}{N}\sum_{t=1}^{N}(u_t-\overline{u})^2 = \sum_{k=1}^{n}(\hat{A}_k^2+\hat{B}_k^2) \quad ,$$

as was to be shown.

A more pedestrian derivation is given by using (A.1.4) of the Appendix to sum

(1.6.3) $$\frac{4}{N^2}\sum_{k=1}^{n}\ \sum_{n=1}^{N}\ \sum_{s=1}^{N} u_r u_s \cos\frac{(r-s)2\pi k}{N} \quad .$$

This becomes, for $r \neq s$,

$$\frac{r}{N^2}\sum_{r}^{N}\sum_{s}^{N}\frac{\cos(\frac{N-1}{2})(r-s)\,\frac{2n}{N} - \cos\,(\frac{n+1}{2})\frac{2\pi}{N}(r-s)-1+\cos(r-s)\frac{2\pi}{N}}{s(1-\cos\frac{(r-s)2\pi k}{N})} \quad ,$$

where all trigonometric terms cancel, leaving the contribution for $r \neq s$ as $-\frac{2}{N^2}\sum_{r}^{N}\sum_{s}^{N} u_r u_s$. Summing over k for $r = s$, each cosine term in (1.6.3) is 1, so that the sum of the n intensities for $2n+1$ odd is

$$\sum_{k=1}^{n}(\hat{A}_k^2+\hat{B}_k^2)= \frac{4}{N^2}\left[(\frac{N-1}{2})\sum_{t=1}^{N} u_t^2 - \frac{1}{2}\sum_{\substack{r\\ r\neq s}}^{N}\sum_{s}^{N} u_r u_s\right]$$

$$= \frac{2}{N}\sum_{t=1}^{N}(u_t-\overline{u})^2 \quad ,$$

as we wished to show.

For N even, so that it is of the form $N = 2(n+1)$, Fisher [5] suggests that only the first n complete harmonics be considered, and states in effect that the sum of these n intensities is

$$(1.6.4) \qquad \frac{2}{N}\sum_{t=1}^{N}(u_t-\bar{u})^2 - \frac{2\left[\Sigma u_r(-)^{r+1}\right]^2}{N^2}$$

Actually, Fisher's result is $\Sigma(u_t-\bar{u})^2 - \frac{\left(\Sigma u_r(-)^{r+1}\right]^2}{N}$,

but the presence of our factor $\frac{1}{N}$ is due to the fact that his esimates of A and B involve $\sqrt{\frac{2}{N}}$, and not $\frac{2}{N}$, as chosen for this study. We shall now find the sum of the n complete intensities.

Starting with (4.2.4), this sum is, (for $\theta = \frac{2\pi(r-s)}{N}$),

$$(1.6.5) \qquad \frac{4}{N^2}\left[\sum_{r}^{N}\sum_{\substack{s\\ r\neq s}}^{N} u_r u_s \left(\frac{\cos\theta(\frac{N-2}{2})-\cos\frac{N\theta}{2} - 1+\cos\theta}{2(1-\cos\theta)}\right)+ n\sum_{t=1}^{N}u_t^2\right],$$

where the first expression is equal to 0 for r-s odd, and equal to

$\frac{-4}{N}\sum_{\substack{r,s\\ r\neq s}} u_r v_s$ for r-s even. (1.6.5) can thus be written as

$$(1.6.6) \qquad \frac{4}{N^2}\left[n\sum_{t=1}^{N}u_t^2 - \frac{1}{2}\left(\sum_{r}^{N}\sum_{\substack{s\\ r\neq s}}^{N} u_r u_s + \sum_{r\neq s}^{N}\sum^{N} u_r u_s(-)^{r-s}\right)\right].$$

Since

$$\left[\sum_{t=1}^{N}u_t(-)^{t+1}\right]^2 = \sum_{t=1}^{N}u_t^2 + \sum_{r\neq s}\sum u_r u_s(-)^{r-s} \qquad ,$$

the above is equal to

$$\frac{4}{N^2}\left[n\sum_{t=1}^{N}u_t^2 - \frac{1}{2}\left[\sum_{r\neq s}\Sigma u_r u_s+\left(\sum_{t=1}^{N}u_t(-)^{t+1}\right)^2 - \sum_{t=1}^{N}u_t^2\right]\right].$$

Since $N = 2(n+1)$, this becomes

$$\frac{4}{4(n+1)^2}\left[n\sum_{t=1}^{N} u_t^2 - \frac{1}{2}\left[\left[\sum_{t=1}^{N} u_t(-)^{t+1}\right]^2 + \left(\sum_{t=1}^{N} u_t\right)^2 - 2\sum_{t=1}^{N} u_t^2\right]\right]$$

$$= \frac{1}{(n+1)}\left[\sum_{t=1}^{N} u_t^2 - \frac{\left(\sum_{t=1}^{N} u_t\right)^2}{2(n+1)} - \frac{\left(\sum_{t=1}^{N} u_t(-)^{t+1}\right)^2}{2(n+1)}\right]$$

$$= \frac{1}{(n+1)}\left[\sum_{t=1}^{N} (u_t-\bar{u})^2 - \frac{\left(\sum_{t=1}^{N} u_t(-)^{t+1}\right)^2}{2(n+1)}\right].$$

To show the similarity to the case $N = 2n+1$, this may be written as

$$= \frac{2}{N}\sum_{t=1}^{N} (u_t-\bar{u})^2 - \frac{2\left(\sum_{t=1}^{N} u_t(-)^{t+1}\right)^2}{N^2},$$

as originally stated in (1.6.4) .

1.7. The Expectation of the Periodogram Ordinate in the Vicinity of the True Period

Let the model be of the usual standard univariate form

$$u_t = \quad + A\cos t\lambda + B\sin t\lambda + \varepsilon_t \quad, \tag{1.7.1}$$

with the assumptions (1.2.2) on the residuals ε_t satisfied. Let us compute the expectation of the intensity $\hat{A}^2 + \hat{B}^2$ for the trial period corresponding to $\theta = \frac{2\pi k}{N}$:

$$E(\hat{A}^2+\hat{B}^2) = \frac{4E}{N^2}\left(\sum_{t=1}^{N} u_t \cos t\theta\right)^2 + \frac{4}{N^2}E\left(\sum_{t=1}^{N} u_t \cos t\theta\right)^2$$

$$+ \frac{4E}{N^2} \left[\sum_{t=1}^{N} u_t (A_1 + A \cos \lambda t + B \sin t\lambda + G_t) \cos t\theta \right]^2$$

(1.7.2)

$$+ \frac{4E}{N^2} \left[\sum_{t=1}^{N} u_t (A_o + A \cos t\lambda + B \sin t\lambda + \varepsilon_t) \sin t\theta \right]^2$$

$$= \frac{4A^2}{N^2} \left(\sum_{t=1}^{N} \cos t\lambda \cos t\theta \right)^2 + \frac{4A^2}{N^2} \left(\sum_{t=1}^{N} \cos t\lambda \sin t\theta \right)^2$$

$$+ \frac{4B^2}{N^2} \left(\sum_{t=1}^{N} \sin t\lambda \cos t\theta \right)^2 + \frac{4B^2}{N^2} \left(\sum_{t=1}^{N} \sin t\lambda \sin t\theta \right)^2$$

$$+ \frac{8AB}{N^2} \left(\sum_{t=1}^{N} \cos t\lambda \cos t\theta \right) \left(\sum_{t=1}^{N} \sin t\lambda \cos t\theta \right)$$

$$+ \frac{8AB}{N^2} \left(\sum_{t=1}^{N} \cos t\lambda \sin t\theta \right) \left(\sum_{t=1}^{N} \sin t\lambda \sin t\theta \right)$$

$$+ \frac{4\sigma^2}{N} .$$

For $\lambda = \frac{2\pi(k + \frac{1}{2})}{N}$, the model period is farthest from any of the possible trial periods of the usual form $\frac{2\pi k}{N}$. Using the trigonometric sums (A.5.1) - (A.5.8) of the Appendix, we may evaluate the expectation (1.7.2) when the true period is half-way between two consecutive sub-multiples of N as

$$(1.7.3)\quad E(\hat{A}^2+\hat{B}^2)= \frac{4\sigma^2}{N} + \frac{4A^2}{N^2} + \frac{4A}{N^2}\left[\frac{-N}{\pi}\frac{4k}{4k+1} \right]^2$$

$$+ \frac{4B^2}{N^2}\left[\frac{N}{\pi}\frac{4k+2}{4k+1} \right]^2 + \frac{8AB}{N^2}\left[-\frac{N}{\pi}\frac{(4k+2)}{(4k+1)} \right].$$

Neglecting terms in N^{-2} and N^{-1}, we are left with

$$(1.7.4)\quad E(A^2+B^2) = \frac{4}{\pi^2}\left[A^2\left[\frac{4k}{4k+1} \right]^2 + B^2\left[\frac{4k+2}{4k+1} \right]^2 \right]$$

$$\approx \frac{4}{\pi^2}(A^2+B^2) \quad .$$

It follows from these computations that the expectation of the periodogram ordinate when the true period is one-half a unit between it and the next ordinate is about 0.406 that of the population intensity A^2+B^2, or 0.406 times the expectation of the ordinate resulting from trial and model periods being exactly equal.

In similar fashion, the expectation of the ordinate for $\lambda = \frac{\pi(4k+1)}{2N}$, may be found from the Appendix sums (A. 6. 1) - (A.6.6). Omitting the details of computation, save to note that the terms in AB are conveniently of the order N^{-2}, and that all the coefficients of A and B are asymptotically equal, the expectation for this value of I is

$$(1.7.5)\qquad E(A^2+B^2) \approx \frac{8}{\pi}(A^2+B^2) = 0.81(A^2+B^2)$$

It is evident that a model period as close as one-quarter of a unit to any trial period would, for large N, be reflected in that trial period's intensity.

These results have been known for some time, the first one having been discussed under slightly different circumstances by Whittle [18], p. 107. They have been presented here not only to illustrate the effect of non-integral sub-multiple periods upon the periodogram, but also because it is felt that more explicit derivations of the summations involved would be useful.

1.8. The Power Function for Walker's Test

The probability of not rejecting the null hypothesis of any simple sinusoidal component when σ^2 is known can be found quite simply. Consider Walker's cumulative distribution (1.4.9):

$$P(x > x') = 1 - \left[1 - e^{-\frac{x'}{2}}\right]^n$$

where in the usual notation, $x = \frac{N(\hat{A}^2 + \hat{B}^2)}{2\sigma^2}$. For a fixed $x' = x_o$ corresponding to a Type I error of prescribed size α, we wish to know the probability

$$\beta = P\left[x < x_o \mid H_a: A^2 + B^2 = K\right] , \tag{1.8.1}$$

that is, the probability of accepting H_o: $A^2 + B^2 = 0$, given that H_a: $A^2 + B^2 = K$ is true. If this alternative is correct, then one of the n x_i has a non-central χ^2 distribution with 2 degrees of freedom and non-centrality parameter $\lambda = \frac{N}{2}(A^2 + B^2)$, while the remaining (n-1) of the x_i have central $\chi^2_{(2)}$ distributions.

Fisher [4], p. 670, gives the probability element for a non-central χ^2, i.e., χ'^2, variate with n degrees of freedom and non-centrality parameter λ as

$$(1.8.2)\quad \phi(\chi'^2,\lambda,n)d\chi'^2 = \frac{e^{-\frac{\chi'^2+\lambda^2}{2}}(\chi'^2)^{\frac{n-2}{2}}}{2^{\frac{n}{2}}} \sum_{j=0}^{\infty} \frac{(\chi'^2)^j \lambda^j d\chi'^2}{\Gamma(j+1)\Gamma(j+\frac{n}{2})r^j} .$$

For n = 2, this reduces to

$$(1.8.3)\quad \phi(\chi'^2, \lambda, 2)d\chi'^2 = \frac{e^{-\frac{\chi'^2+\lambda}{2}}}{2} \sum_{j=0}^{\infty} \frac{(\chi'^2)^j \lambda^j d\chi'^2}{(j!)^2 4^j} .$$

The probability of accepting H_o, given H_a true, is then

$$\beta = 1 - \left\{ 1 - [1-e^{\frac{-x_o}{2}}]^{n-1} \int_0^{x_o} \phi(\chi'^2,\lambda,2)d\chi'^2 \right\}$$

$$(1.8.4)\quad = [1-e^{\frac{-x_o}{2}}]^{n-1} \int_0^{x_o} \frac{e^{-\frac{\chi'^2+\lambda}{2}}}{2} \sum_{j=0}^{\infty} \frac{(\chi'^2)^j \lambda^j d\chi'^2}{(j!)^2 4^j} .$$

The power function becomes

$1-\beta =$

$$1-[1-e^{-\frac{x_o}{2}}]^{n-1}\ \frac{e^{-\frac{\lambda}{2}}}{2} \int_0^{x_o} e^{-\frac{\chi'^2}{2}} \sum_{j=0}^{\infty} \frac{(\chi'^2)^j \lambda^j}{(j!)^2 2^{2j}}\, d\chi'^2 \quad . \tag{1.8.5}$$

No simpler expression appears to be available at this time. We may put a very loose upper bound on (1.8.5) by interchanging the integral and summation signs, so that the power function may be written as

$$1-[1-e^{-\frac{x_o}{2}}]^{n-1}\ \frac{e^{-\frac{\lambda}{2}}}{2} \sum_{j=0}^{\infty} \frac{\lambda^j}{(j!)^2 4^j} \int_0^{x_o} e^{-\frac{\chi'^2}{2}} (\chi'^2)^j d\chi'^2 \tag{1.8.6}$$

$$=1-[1-e^{-\frac{x_o}{2}}]^{n-1} e^{-\frac{\lambda}{2}} [1-e^{-\frac{x_o}{2}}] + \frac{1}{2} \sum_{j=1}^{\infty} \frac{\lambda^j}{(j!)^2 4^j} \int_0^{x_o} e^{-\frac{u}{2}} u^j du \quad .$$

This is justified by the uniform convergence of the series $\sum_{j=1}^{\infty} \frac{\lambda^j}{(j!)^2}$,

That such convergence holds is obvious if we note that

$\sum_{j=1}^{\infty} \frac{\lambda^j}{(j!)^2} \leq \frac{\lambda^j}{j!}, \sum_{j=1}^{\infty} \frac{\lambda^j}{j!} = e^{\lambda}$, so that Weierstrass's M-test is satisfied.

Since each of the integrals $\int_0^{x_o} e^{-\frac{u}{2}} u^j du$ is positive, it is evident that the power function is bounded above by

$$1-[1-e^{-\frac{x_o}{2}}]^{n} e^{-\frac{\lambda}{2}} \quad . \tag{1.8.7}$$

For j a positive integer, the integral may be integrated by parts repeatedly, so that

$$(1.8.8) \qquad \int_0^{x_o} u^j e^{-\frac{u}{2}} du = 2^{j+1} j! \left[1 - e^{-\frac{x_o}{2}} \sum_{i=o}^{j} \frac{1}{(j-i)!} \left(\frac{x_o}{2} \right)^{j-i} \right].$$

Employing this in (1.8.6), the power function becomes

$$(1.8.9) \quad 1-(1-e^{-\frac{x_o}{2}})^{n-1} e^{-\frac{\lambda}{2}} \sum_{j=o}^{\infty} \frac{\lambda^j}{j!2^j} \left[1-e^{-\frac{x_o}{2}} \sum_{i=o}^{j} \frac{1}{(j-i)!} \left(\frac{x_o}{2} \right)^{j-i} \right].$$

We shall now demonstrate that Walker's test is unbiassed, i.e., that the power is greater than the probability of the first kind of error, α. We thus wish to show that

$$(1.8.10) \quad \begin{aligned} & 1 - (1-e^{-\frac{x_o}{2}})^n \leq \\ & 1-(1-e^{-\frac{x_o}{2}})^{n-1} e^{-\frac{\lambda}{2}} \sum_{j=o}^{\infty} \frac{\lambda^j}{j!2^j} \left[1-e^{-\frac{x_o}{2}} \sum_{j=o}^{j} \frac{1}{(j-i)!} \left(\frac{x_o}{2} \right)^{j-i} \right]. \end{aligned}$$

This simplifies to

$$e^{-\frac{x_o}{2}} \leq e^{-\frac{\lambda+x_o}{2}} \sum_{j=o}^{\infty} \sum_{i=o}^{j} \frac{\lambda^j x_o^{j-i}}{j!2^j (j-i)!} ,$$

or

$$0 \leq e^{-\frac{\lambda}{2}} \sum_{j=o}^{\infty} \frac{\lambda^j}{j!2^j} \left(\frac{x_o}{2} + \dots + \frac{x_o^j}{j!2^j} \right) .$$

Since $x_0, \lambda \geq 0$, it is readily seen that the inequality holds. Hence, Walker's test is unbiassed.

CHAPTER II

A TEST FOR SIGNIFICANCE OF PERIODIC COMPONENTS IN BIVARIATE TIME SERIES

2.1. Introduction

We shall now consider the case of two parallel series of observations ordered in time, and each with a single sinusoidal component. In the case of the same model period in each series, an exact test for the significance of a trial period is obtained. This test is based upon the determinant of the matrix of intensities and cross-intensities of the two series.

The motivation for such a model and test might for example be provided from astronomy by a series of simultaneous observations on the brightness and radial velocity of a variable star.

2.2 The Bivariate Model

The general bivariate model used for the results of this chapter is

$$u_t = C_1 + A_1 \cos \frac{2\pi k}{N} t + B_1 \sin \frac{2\pi k}{N} t + \varepsilon_t \quad .$$

(2.2.1)

$$v_t = C_2 + A_2 \cos \frac{2\pi m}{N} t + B_2 \sin \frac{2\pi m}{N} t + \delta_t \quad .$$

t runs through the integers 1, 2, ..., N, N being restricted to be odd to take advantage of certain properties of sums of intensities of a complete Fourier sequence. k and m are integers from the set 1, 2, ..., n, where, as in the univariate model, n will be used henceforth to denote $\frac{N-1}{2}$. The test of significance developed in this chapter will be based upon $k \equiv m$ to avoid the introduction of nuisance parameters, although the effects of different model periods, as well as ones not of the integral sub-multiple form will be considered in § 2.3 and Chapter III.

The following assumptions on the errors will be made:

$$E(\varepsilon_t) = E(\delta_t) = 0 \quad ,$$

$$(2.2.2) \qquad E(\varepsilon_r, \varepsilon_s) = \sigma_1^2 \delta_{rs}, \; E(\delta_r, \delta_s) = \sigma_2^2 \delta_{rs} \quad ,$$

$$\mathrm{Cov}(\varepsilon_r, \delta_s) = \delta_{rs} \, \rho \, \sigma_1 \sigma_2 \quad ,$$

where δ_{rs} is the Kronecker δ, assuming the value 1 for r equal to s, and the value 0 for r and s different. A more general scheme of correlation will be assumed in Chapter III to determine its effect upon certain expected values. We assume that $\rho^2 < 1$; this is necessary and sufficient for the covariance matrix to be positive definite, and therefore for the consistency of the assumptions (2.2.2).

For the test of significance obtained in this chapter, it is further necessary to assume that the $(\varepsilon_t, \delta_t)$ have a bivariate normal distribution with means and covariance matrix determined by (2.2.2).

2.3. Estimation of the Parameters C_1, C_2, A_1, B_1, A_2, B_2

We shall begin this section with a theorem on multivariate least squares. For the case of equal periods in each of the series of (2.2.1), this theorem states that the estimates of C_1, A_1, B_1, and

C_2, A_2, B_2 are the same as those given by (1.3.5) for a single series.

First define the following matrices for a general multivariate regression model:

Matrix of Predictors:

$$(2.3.1) \qquad X(N \times p) = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ x_{N1} & \cdots & x_{Np} \end{bmatrix} = \begin{bmatrix} \underline{x}_1' \\ \cdot \\ \cdot \\ \cdot \\ \underline{x}_N' \end{bmatrix} .$$

Multiple Predictand Matrix:

$$(2.3.2) \qquad Y(N \times q) = \begin{bmatrix} y_{11} & \cdots & y_{1q} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ y_{N1} & \cdots & y_{Nq} \end{bmatrix} = \begin{bmatrix} \underline{y}_1' \\ \cdot \\ \cdot \\ \cdot \\ \underline{y}_N' \end{bmatrix} .$$

Regression Coefficients:

$$(2.3.3) \qquad B(p \times q) = \begin{bmatrix} \beta_{11} & \cdots & \beta_{1q} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \beta_{p1} & \cdots & \beta_{pq} \end{bmatrix} = (\underline{\beta}_1, \ldots, \underline{\beta}_q) .$$

Errors:

$$(2.3.4) \qquad \Delta(N \times q) = \begin{bmatrix} e_{11} & \cdots & e_{1q} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ e_{N1} & \cdots & e_{Nq} \end{bmatrix} = \begin{bmatrix} \underline{e}_1' \\ \cdot \\ \cdot \\ \cdot \\ \underline{e}_N' \end{bmatrix} .$$

The elements e_{ij} of Δ are random variables with

$$(2.3.5)\qquad \begin{aligned} &E(e_{ij}) = 0, \ (i = 1, \ldots, N; \ j = 1, \ldots, q) \ , \\ &E(\Delta\Delta') = \Lambda \ (q \times q), \text{ non-singular.} \end{aligned}$$

For each value of i, $(i = 1, \ldots, N)$, the elements e_{ij} of the row vector $\underline{e}_i'$ have the multivariate normal distribution $N [\underline{0}'(q \times 1), \Lambda \ (q \times q)]$. These N distributions are independent, so that $E(e_{rm} \ e_{sn}) = \delta_{rs} \rho_{mn} \sigma_m \sigma_n$, where δ_{rs} is the Kronecker symbol described in Section 2.2.

We shall now prove the following theorem:

Theorem:

For the multiple predictand regression model

$$(2.3.6)\qquad Y(N \times q) = X(N \times p) \ B(p \times q) + \Delta(N \times q) \ ,$$

the pq estimates b_{ij} of the elements of B are given p at a time by the least squares solutions for the q predictands taken separately.

Proof:

From the conditions (2.3.5), the likelihood of the matrix Δ may be written as

$$(2.3.7)\qquad L(\Delta) = \text{const. } \exp \Big[- \frac{1}{2} \sum_{t=1}^{N} \underline{e}_t' \Lambda^{-1} \underline{e}_t \Big]$$

$$= \text{const. } \exp \left[-\frac{1}{2} \sum_{t=1}^{N} \sum_{k=1}^{q} \sum_{\ell=1}^{q} \lambda^{k\ell} e_{tk} e_{t\ell} \right] .$$

Taking logarithms and substituting for the e_t' from the multiple predictand model (2.3.6), $L(\Delta)$ is a monotonic decreasing function of the exponent:

$$(2.3.8) \quad L^*(\Delta) = \sum_{t=1}^{N} \sum_{k=1}^{q} \sum_{\ell=1}^{q} \lambda^{k\ell} \left[y_{tk} - \sum_{i=1}^{p} \beta_{ik} x_{ti} \right] \left[y_{t\ell} - \sum_{j=1}^{p} \beta_{j\ell} x_{tj} \right].$$

Differentiation with respect to the β_{jk}, with the resulting derivatives set equal to 0, yields

$$(2.3.9) \quad \frac{\partial L^*(\Delta)}{\partial \beta_{jk}} = \sum_{t=1}^{N} \sum_{\ell=1}^{q} \lambda^{k\ell} x_{ti} (y_{tq} - \sum_{i=1}^{p} \beta_{j\ell} x_{tj}) = 0, \quad \begin{matrix} (j=1,\ldots,p). \\ (k=1,\ldots,q). \end{matrix}$$

Next interchange the summations over t and ℓ:

$$(2.3.10) \qquad \sum_{\ell=1}^{q} \lambda^{k\ell} \sum_{t=1}^{N} (x_{ti} y_{t\ell} - \sum_{j=1}^{p} x_{ti} x_{tj} \beta_{j\ell}) = 0 \quad .$$

If we adopt the notation of single-predictand least squares theory, where

$$(2.3.11) \qquad \begin{aligned} a &= X'(P \times N)\, X\, (N \times p) \\ &= \left[\sum_{t=1}^{N} x_{ti} \, x_{tj} \right], \qquad a \neq 0 \end{aligned}$$

$$\underline{g}_{\ell} = X'(p \times N)\,\underline{y}(N \times 1)$$

$$= \left[\sum_{t=1}^{N} x_{ti}\, y_{t\ell} \right] ,$$

we finally obtain

$$(2.3.12) \qquad \sum_{\ell=1}^{q} \lambda^{k\ell} \left(a\underline{\beta}_{\ell} - \underline{g}_{\ell} \right) = 0 \quad .$$

In this set of q equations let $\gamma_{\ell} = a\underline{\beta}_{\ell} - \underline{g}_{\ell}$. Then, since the covariance matrix Λ is non-singular, we have a set of homogeneous linear equations in the γ_{ℓ} whose only solution is given by $\gamma_{\ell} = 0$, ($\ell = 1, \ldots, q$), or,

$$(2.3.13) \qquad \hat{\underline{\beta}}_{\ell} = a^{-1}\underline{g}_{\ell}$$

Thus, the estimates of the elements of B are given by the q matrix equations, each containing p estimates, as was to be shown.

From this theorem it follows that the estimates of C_1, A_1, B_1, C_2, A_2, B_2, for <u>equal</u> model periods are the same as those given by Chapter I, (1.3.5), or in the notation of the bivariate model,

$$(2.3.14) \qquad \begin{aligned} \hat{C}_1 &= \frac{1}{N} \sum_{t=1}^{N} u_t \ , & \hat{C}_2 &= \frac{1}{N} \sum_{t=1}^{N} v_t \ , \\ \hat{A}_1 &= \frac{2}{N} \sum_{t=1}^{N} u_t \cos \frac{2\pi k}{N} t \ , & \hat{A}_2 &= \frac{2}{N} \sum_{t=1}^{N} v_t \cos \frac{2\pi k}{N} t \ , \\ \hat{B}_1 &= \frac{2}{N} \sum_{t=1}^{N} u_t \sin \frac{2\pi k}{N} t \ , & \hat{B}_2 &= \frac{2}{N} \sum_{t=1}^{N} v_t \sin \frac{2\pi k}{N} t \ . \end{aligned}$$

For the case of different model periods, the X matrix (2.3.1) is not the same for each series, so that the theorem does not apply. The resulting estimates are rather interesting, and will now be obtained. We may write the likelihood (2.3.7) as

$$\text{(2.3.15)} \qquad L = \text{const.}\ \exp \left[-\frac{1}{2} \sum_{t=1}^{N} \underline{e}_t' \ \Sigma^{-1} \underline{e}_t \right] ,$$

where

$$e_t = \begin{bmatrix} u_t - C_1 - A_1 \cos \lambda_1 t - B_1 \sin \lambda_1 t \\ v_t - C_2 - A_2 \cos \lambda_2 t - B_2 \sin \lambda_2 t \end{bmatrix} ,$$

$$\lambda_1 = \frac{2\pi k}{N} , \lambda_2 = \frac{2\pi m}{N} , \text{ and } \Sigma^{-1} = \frac{1}{\sigma_1^2 \sigma_2^2 (1-\rho^2)} \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1 p_2 & \sigma_1^2 \end{bmatrix} .$$

Taking logarithms and expanding the summation over r and s, the likelihood (2.3.21) is proportional to

$$\text{(2.3.16)} \qquad \begin{aligned} L' = {} & \sigma_2^2 \sum_{t=1}^{N} (u_t - C_1 - A_1 \cos \lambda_1 t - B_1 \sin \lambda_1 t)^2 \\ & + \sigma_1^2 \sum_{t=1}^{N} (v_t - C_2 - A_2 \cos \lambda_2 t - B_2 \sin \lambda_2 t)^2 \\ & - 2\, \sigma_1 \sigma_2 \sum_{t=1}^{N} (u_t - C_1 - A_1 \cos\lambda_1 t - B_1 \sin\lambda_1 t)(v_t - C_2 - A_2 \cos\lambda_2 t - B_2 \sin\lambda_2 t) . \end{aligned}$$

Differentiating with respect to C_1, C_2, A_1, B_1, A_2, and B_2, using the sums (A.2.1) - (A.2.6) of the Appendix, and setting the resulting expressions equal to 0, the estimates are:

$$\hat{C}_1 = \frac{1}{N} \sum_{t=1}^{N} u_t \quad .$$

$$\hat{C}_2 = \frac{1}{N} \sum_{t=1}^{N} v_t \quad .$$

(2.3.17)

$$\hat{A}_1 = \frac{2}{N} \sum_{t=1}^{N} u_t \cos \lambda_1 t - \rho \frac{\sigma_1}{\sigma_2} \frac{2}{N} \sum_{t=1}^{N} v_t \cos\lambda_1 t \quad .$$

$$\hat{B}_2 = \frac{2}{N} \sum_{t=1}^{N} u_t \sin\lambda_1 t - \rho \frac{\sigma_1}{\sigma_2} \frac{2}{N} \sum_{t=1}^{N} v_t \sin \lambda_1 t \quad .$$

$$\hat{A}_2 = \frac{2}{N} \sum_{t=1}^{N} v_t \cos \lambda_2 t - \rho \frac{\sigma_2}{\sigma_1} \frac{2}{N} \sum_{t=1}^{N} u_t \cos\lambda_2 t \quad .$$

$$\hat{B}_2 = \frac{2}{N} \sum_{t=1}^{N} v_t \sin \lambda_2 t - \rho \frac{\sigma_2}{\sigma_1} \frac{2}{N} \sum_{t=1}^{N} u_t \sin\lambda_2 t \quad .$$

It is especially noteworthy that the assumption of unequal periods has introduced nuisance parameters $\rho\frac{\sigma_2}{\sigma_1}$ and $\rho\frac{\sigma_1}{\sigma_2}$ into the estimates. For this reason, the test proposed in this chapter will be concerned only with equal model periods. Some properties of the more general model will be treated in Chapter III.

2.4. A Bivariate Test for a Simple Sinusoid

From the complete sample Fourier sequence of estimates given by (2.3.18)-(2.3.20), form the 2n intensities

$$I_{1k} = \hat{A}_{1k}^2 + \hat{B}_{1k}^2$$
$$I_{2k} = \hat{A}_{2k}^2 + \hat{B}_{2k}^2 \quad ,$$

and the n "cross intensities"

$$I_{12k} = \hat{A}_{1k}\hat{A}_{2k}+\hat{B}_{1k}\hat{B}_{2k} \quad .$$

These form a matrix whose determinant is

$$(2.4.1) \qquad S_k = \begin{vmatrix} I_{1k} & I_{12k} \\ I_{12k} & I_{2k} \end{vmatrix} = (\hat{A}_{1k}\hat{B}_{2k}-\hat{A}_{2k}\hat{B}_{1k})^2 \quad .$$

It has long been known [3,p.71], from Parseval's Theorem that terms of the form $\hat{A}_{1k}\hat{A}_{2k}+\hat{B}_{1k}\hat{B}_{2k}$ measure the covariance between the two series of the model (2.2.1). In fact, under this model, and the usual error scheme (2.2.4),

$$(2.4.2) \qquad E(\hat{A}_{1k}\hat{A}_{2k}+\hat{B}_{1k}\hat{B}_{2k}) = \frac{4}{N}\rho\sigma_1\sigma_2 + (A_{1k}A_{2k}+B_{1k}B_{2k})\,\delta_{km} \quad ,$$

where δ_{km} is the usual Kronecker symbol, assuming the value 1 for $k = m$, that is, for trial and model periods equal and corresponding to arguments of the form $\frac{2\pi k}{N}t$, and 0 for other trial periods in the orthogonal set. Thus, since the matrix of S_k is a covariance matrix associated with the contribution of the kth trial period, we shall take the determinant S_k as the bivariate analogue of the univariate periodogram intensity. In accordance with the univariate procedure, we now wish the distribution of the greatest of the n intensity determinants, or better, the distribution of some "studentized" function of them.

To obtain this distribution, consider Wilks' [18] distribution of the determinant ξ of the matrix of product-sums of bivariate normal variates:

$$(2.4.3) \qquad F(\xi)d\xi = \frac{2^{n-2}A^{\frac{n-1}{2}}\xi^{\frac{n-2}{2}}e^{-2(A\xi)^{\frac{1}{2}}}d\xi}{(n-2)!}, \quad \xi \geq 0 ,$$

where

$$A = \frac{1}{2^2\sigma_1^2\sigma_2^2(1-\rho^2)} .$$

Wilks' original notation in the 1932 paper involved the sample size N, which has been replaced here by the number of degrees of freedom, $n = N-1$. Wilks' A also contains the factor N^n, since his ξ is a determinant of <u>mean</u> squares, while our intensity determinant is one of product-sums. For the intensity determinant, A thus becomes $[4 \, Var(A_1)Var(A_2)(1-\rho^2)]^{-1}$. Henceforth, Σ will denote the population covariance determinant $\frac{4\sigma_1^2\sigma_2^2(1-\rho^2)}{N^2}$. For $n = 2$ degrees of freedom, the distribution of a sample intensity determinant for no periodic component and errors satisfying the conditions (2.2.2) is

$$(2.4.5) \qquad G(S \quad)dS = \frac{S^{-\frac{1}{2}}e^{-\frac{S^{1/2}}{\Sigma^{1/2}}}dS}{2\Sigma^{1/2}} .$$

If we let $s = 2\Sigma^{1/2}S^{1/2}$, the distribution of s is

$$(2.4.6) \qquad f(s)ds = \frac{1}{2\Sigma} e^{-\frac{S}{2\Sigma}} ds \quad ,$$

or merely a $\chi^2\sigma^2$ distribution with 2 degrees of freedom. Hence Fisher's g becomes

$$(2.4.7) \qquad g = \frac{\max_k \sqrt{S_k}}{\sum_{k=1}^{n} \sqrt{S_k}} \quad ,$$

and has the cumulative distribution (1.5.2) of Chapter I. Thus, using the intensity determinant as a test statistic, we are provided with an exact significance test for a periodic component in two parallel time series.

The sum $\sum_{k=1}^{n} \sqrt{S_k}$ of the denominator of g can be written as

$$(2.4.8) \qquad \sum_{k=1}^{n} \left| \sum_{r=1}^{N} \sum_{s=1}^{N} u_r v_s \sin \frac{2\pi k(r-s)}{N} \right| \ .$$

Because the positive square root is taken in each case, no exact expression is available at this time for the sum, nor has any close practical upper bound been obtained for it. Examination for small values of N has indicated that its value may involve irregular sums of cross products of the u_r, v_s, whose terms are multiplied by the often irrational value of the sine factor involved. Even for positive observations, it does not appear that relations similar to those for the sums of univariate intensities exist.

We may obtain a lower bound on the sum $\sum_{k=1}^{n} | S_k^{1/2} |$ in the following manner. Disregard the absolute value signs, and write the sum as

$$(2.4.9) \qquad \frac{4}{N^2} \sum_{k=1}^{n} \sum_{r=1}^{N} \sum_{s=1}^{N} u_r v_s \sin \frac{2\pi(r-s)k}{N} .$$

Using the fundamental sum (A.1.5) of the Appendix, with the appropriate values of a, b, θ, and ρ inserted, this becomes

$$(2.4.10) \qquad \frac{4}{N^2} \sum_{\substack{r \\ r \neq s}}^{N} \sum_{s}^{N} u_r v_s \left[\frac{\sin\frac{(N-1)2\pi(r-s)}{2N} - \sin\frac{(N+1)2\pi(r-s)}{2N}}{2(1-\cos \frac{2\pi(r-s)}{N})} + \frac{\sin \frac{2\pi(r-s)}{N}}{2(1-\cos \frac{2\pi(r-s)}{N})} \right] .$$

With the aid of some trigonometric identities, this simplifies further to

$$(2.4.11) \qquad \frac{4}{N^2} \sum_{\substack{r \\ r \neq s}}^{N} \sum_{s}^{N} u_r v_s \frac{\sin \frac{\pi(r-s)}{N} \left[\cos \frac{\pi(r-s)}{N} - (-)^{r-s} \right]}{2(1-\cos^2 \frac{\pi(r-s)}{N})}$$

$$= \frac{2}{N^2} \sum_{\substack{r \\ r \neq s}}^{N} \sum_{s}^{N} u_r v_s \left[\frac{\cos \frac{\pi(r-s)}{N} - (-)^{r-s}}{\sin \frac{\pi(r-s)}{N}} \right] .$$

Since it is well known that

$$|a_1+a_2+a_3+\ldots+a_n| \le |a_1| + |a_2| + \ldots + |a_n| \quad ,$$

we have the lower bound

$$(2.4.12) \quad \frac{2}{N^2}\left| \sum_{r}^{N} \sum_{\substack{s \\ r \neq s}}^{N} u_r v_s \left[\frac{\cos\frac{\pi(r-s)}{N} - (-)^{r-s}}{\sin\frac{\pi(r-s)}{N}} \right] \right| \le \sum_{k=1}^{n} \left| S_k^{\frac{1}{2}} \right| \quad .$$

Although Chapter III is devoted completely to the expectation of the intensity determinant under a variety of assumptions, we shall give some basic results here as a justification of its use as the test statistic. Under the basic bivariate model (2.2.1), with equal periods, the expectation for model period exactly equal to the trial period can be found from the general expression (3.2.3) of Chapter III to be

$$(2.4.13) \quad E(\hat{A}_1\hat{B}_2-\hat{A}_2\hat{B}_1)^2 = (A_1B_2-A_2B_1)^2 + \frac{2}{N}\left[\sigma_1^2(A_1^2+B_1^2)+\sigma_2^2(A_2^2+B_2^2)\right]$$

$$- \frac{4\rho\sigma_1\sigma_2}{N}(A_1A_2+B_1B_2) + \frac{8\sigma_1^2\sigma_2^2}{N^2}(1-\rho^2) \quad .$$

Under the null hypothesis of no periodic component, that is, for A_1, B_1, A_2, B_2 all equal to zero, the determinant's expectation is simply

$$(2.4.14) \quad \frac{8\sigma_1^2\sigma_2^2}{N^2}(1-\rho^2) \quad ,$$

as may be checked by direct integration for the first moment of the distribution of S_k in (2.4.5). The variance under the null hypothesis is best found by recourse to the distribution function (2.4.5) as well, and is

$$(2.4.15) \qquad 320 \frac{\sigma_1^4 \sigma_2^4}{N^4} (1- \rho^2)^2 \qquad .$$

From the first expectation above, one serious objection to the use of the intensity determinant as a test statistic is immediately evident. Unlike the univariate intensity, the part of the expectation free of terms in N^{-1} can vanish if

$$(2.4.16) \qquad \frac{A_1}{B_1} = \frac{A_2}{B_2}$$

Further, the complete absence of a periodic component in one series could obscure the presence of a real one in the second series.

2.5. Other Possible Statistics

The deficiency mentioned in the last paragraph of the previous Section might possibly be alleviated by considering some functions of the roots of the determinantal equation

$$(2.5.1) \qquad \left| S_k - \frac{2\lambda}{N} s_k^* \right| = 0 \qquad ,$$

where s_k^* is an independent matrix of the error sums of squares and cross products obtained after the contribution due to a given kth sinusoid has been removed. It is known that S_k and s_k^* are independently distributed. Or, the roots of the equation

$$(2.5.2) \qquad |S_k - \theta(S_1 + S_2 \ldots + S_n)| = 0$$

might be used. Actually, there is a correspondence between these two equations, since $s_k^* = \frac{N}{2}(S_1+S_2+ \ldots S_n - S_k)$, so that the roots λ_{sk} of the first may be expressed as a function of those of the second by means of the relation

$$(2.5.3) \qquad \lambda_{ik} = \frac{\theta_{ki}}{1-\theta_{ki}} \quad .$$

One possible statistic might be the sum of the roots of (2.5.1). The distribution of this sum was obtained by Hotelling [9] as an incomplete Beta integral, and the sum was shown to be free from the effect of collinearity that causes the intensity determinant to be zero when a real period may be present. However, the distribution of $\max_k (\lambda_{1k}+\lambda_{2k})$ has not been obtained, and since s_k^* changes for each intensity determinant S_k, the problem is further complicated. Because the determinant of a sum of matrices is not necessarily equal to the sum of the determinants, the use of Cochran's transformations [2] to derive Fisher's distribution is prohibited.

The product of the roots of the equation (2.5.1) might conceivably form a test criterion. Since this product is equal to the determinant of S divided by that of s^*, it is quite similar to the mere intensity determinant, and has been shown by Wilks [18] to have the same regrettable property of vanishing whenever all of the observations lie on a straight line. Again, the same problems mentioned in the previous paragraph would arise in attempting to find the distribution of the maximum product of roots, although unlike the sum, the distribution of a single product is an elementary closed function readily transformable into the variance-ratio distribution.

The trace of the intensity matrix S_k seems to be ruled out for general use. Even if both series were measured in the same units, it seems likely that the variances and covariance would enter as nuisance parameters in such a way as to resist "studentisation" by dividing by, say, the sum of the traces for the complete Fourier sequence.

Many other possible statistics might be investigated, notably the greatest or least roots of the second equation, $|S-\theta(S_1+ \ldots S_k)| = 0$. This equation has a slightly greater intuitive appeal over (2.5.1), since it is so formally analogous to Fisher's g ratio. As for the sum of the roots, the distribution of the greatest root involves incomplete integrals.

In closing this discussion of other statistics than the intensity determinant, it would be interesting to note a property of that quantity that makes the need of alternative statistics imperative if one wishes to extend these methods beyond the bivariate case. For the general p-series problem, the intensity matrix becomes

$$(2.5.4)\qquad \begin{bmatrix} A_1^2+B_1^2 & \cdot & \cdot & \cdot & A_1A_p+B_1B_p \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ A_1A_p+B_1B_p & \cdot & \cdot & \cdot & A_p^2+B_p \end{bmatrix}$$

With a little insight, this may be factorized into the product of a (p x 2) matrix and its transpose:

$$(2.5.5)\qquad \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ A_p & B_p \end{bmatrix} \begin{bmatrix} A_1 & A_2 & \cdot & \cdot & \cdot & A_p \\ B_1 & B_2 & \cdot & \cdot & \cdot & B_p \end{bmatrix} .$$

Hence, the rank of any intensity matrix is at most 2, so that the use of its determinant as a test statistic is meaningless.

CHAPTER III

CERTAIN EXPECTATIONS OF THE INTENSITY DETERMINANT

3.1. Introduction

It is the purpose of this chapter to obtain the expectation of the intensity determinant under the alternative hypothesis of Chapter II, (2.2.1), and for different estimates of the A_1, A_2, B_1, B_2. It has been necessary to restrict nearly all of the results to the point at which the trial and model periods are equal, and then to restrict the angular argument to be of the form $2\pi k/N$, $k = 1, 2, \ldots, (N-1)/2$. Otherwise, the trigonometric sums would be retractable.

3.2. The Intensity Determinant under the Model of Chapter II

We assume that the parallel series may be represented by the basic model of Chapter II, that is,

$$u_t = A_{o1}+A_1\cos \lambda_1 t+B_1\sin \lambda_1 t + \varepsilon_t \quad ,$$

$$(3.2.1) \qquad v_t = A_{o2}+A_2 \cos \lambda_2 t+B_2\sin \lambda_2 t + \delta_t \quad .$$

At this point no assumptions are made on λ_1, λ_2, ε_t, and δ_t, except that the estimates of A_1, A_2, B_1, B_2 are assumed to have been made for $\lambda_1 = \lambda_2$, with the period a sub-multiple of the number of observations in the series. Under this representation of u_t, v_t, the intensity determinant may be expanded as

$$(3.2.2) \qquad [\hat{A}_1\hat{B}_2 - \hat{A}_2\hat{B}_1]^2$$

$$\frac{16}{N^4}\Big[(\Sigma(A_{o1}+A_1\cos\lambda_1 t\cos\theta t+B_1\sin\lambda_1 t\cos\theta t + \varepsilon_t\cos\theta t)) \times$$
$$\times(\Sigma(A_{o2}+A_2\cos\lambda_2 t\sin\theta t + B_2\sin\lambda_2 t\sin\theta t + \delta_t\sin\theta t)$$
$$-(\Sigma(A_{o2}+A_2\cos\lambda_2 t\cos\theta t + B_2\sin\lambda_2 t\cos\theta t + \delta_t\cos\theta t))$$
$$\times(\Sigma(A_{o1}+A_1\cos\lambda_1 t\sin\theta t + B_1\sin\lambda_1 t\sin\theta t + \varepsilon_t\sin\theta t))\Big]^2$$

The summations here, and elsewhere in this chapter, are all over the values 1, ..., N of t. For the trial argument θ of the form $2\pi k/N$, the terms involving A_{o1}, A_{o2}, vanish, and for other values of θ, they are of order small enough to be neglected. If we make that elimination, distribute the summation operators, and carry out multiplication between the quantities in parentheses, this may be written as

$$(3.2.3) \quad [\hat{A}_1\hat{B}_2 - \hat{A}_2\hat{B}_1]^2 =$$

$$\frac{16}{N^4}\Big[\Big\{A_1A_2(\Sigma\cos\theta t\cos\lambda_1 t)(\Sigma\cos\lambda_2 t\sin\theta t)+A_1B_2(\Sigma\cos\lambda_1 t\cos\theta t)(\Sigma\sin\lambda_2 t\sin\theta t)$$
$$+B_1A_2(\Sigma\sin\lambda_1 t\cos\theta t)(\Sigma\cos\lambda_2 t\sin\theta t)+B_1B_2(\Sigma\sin\lambda_1 t\cos\theta t)(\Sigma\sin\lambda_2 t\sin\theta t)$$
$$-A_2A_1(\Sigma\cos\theta t\cos\lambda_2 t)(\Sigma\cos\lambda_1 t\sin\theta t)-A_2B_1(\Sigma\cos\lambda_2 t\cos\theta t)(\Sigma\sin\lambda_1 t\sin\theta t)$$
$$-B_2A_1(\Sigma\sin\lambda_2 t\cos\theta t)(\Sigma A_1\cos\lambda_1 t\sin\theta t)-B_1B_2(\Sigma\sin\lambda_2 t\cos\theta t)(\Sigma\sin\lambda_1 t\sin\theta t)\Big\}$$
$$+\Big\{(A_1\Sigma\cos\theta t\cos\lambda_1 t+B_1\Sigma\sin\lambda_1 t\cos\theta t)\Sigma\delta_t\sin\theta t$$
$$+(A_2\Sigma\cos\lambda_2 t\sin\theta t+B_2\Sigma\sin\lambda_2 t\sin\theta t)\Sigma\varepsilon_t\cos\theta t$$

$$-(A_2\Sigma\cos\theta t\cos\lambda_2 t+B_2\Sigma\sin\lambda_2 t\cos\theta t)\Sigma\varepsilon_t\sin\theta t$$

$$-(A_1\Sigma\cos\lambda_1 t\sin\theta t+B_1\Sigma\sin\lambda_1 t\sin\theta t)\Sigma\delta_t\cos\theta t\Big\}$$

$$+\Big\{(\Sigma\varepsilon_t\cos\theta t)(\Sigma\delta_t\sin\theta t)-(\Sigma\delta_t\cos\theta t)(\Sigma\varepsilon_t\sin\theta t)\Big\}\quad\Big]^2$$

We shall denote the three braced terms as I, II, and III.

The expectation of the intensity determinant under the assumptions (2.2.2) of uncorrelated residuals of zero expectation may now be obtained. Expanding the trinomial in I, II, III, one obtains the individual expectations:

$$E(I^2) = I^2 \quad ,$$

(3.2.4) $E(I.II) = 0$, since $E\varepsilon_t = E\delta_t = 0$;

$E(II.III) = 0$, since all expectations are either μ_{111}, μ_{21} or μ_{12} ;

$E(I.III) = 0$; since $\sum_{t=1}^{N} \sin\theta t \cos\theta t = 0$.

$$EIII^2 = \frac{\sigma_1^2\sigma_2^2(1-\rho^2)N^2}{2} \quad ,$$

$$EII^2 = \frac{N\sigma_2^2}{2}\Big[\Big\{A_1\Sigma\cos\theta t\cos\lambda_1 t+B_1\Sigma\sin\lambda_1 t\cos\theta t\Big\}^2$$

$$+\Big\{A_1\Sigma\cos\lambda_1 t\sin\theta t+B_1\Sigma\sin\lambda_1 t\sin\theta t\Big\}^2\Big]$$

$$+\frac{N\sigma_1^2}{2}\Big[\Big\{A_2\Sigma\cos\lambda_2 t\sin\theta t+B_2\Sigma\sin\lambda_2 t\sin\theta t\Big\}^2$$

$$+\Big\{A_2\Sigma\cos\theta t\cos\lambda_2 t+B_2\Sigma\sin\lambda_2 t\cos\theta t\Big\}^2\Big]$$

$$-N\rho\sigma_1\sigma_2 \left[\left\{ A_1\Sigma\cos\theta t\cos\lambda_1 t + B_1\Sigma\sin\lambda_1 t\cos\theta t \right\} \right.$$

$$A_2\Sigma\cos\theta t\cos\lambda_2 t + B_2\Sigma\sin\lambda_2 t\cos\theta t$$

$$+ \quad A_2\Sigma\cos\lambda_2 t\sin\theta t + B_2\Sigma\sin\lambda_2 t\sin\theta t$$

$$\left. \times \quad A_1\Sigma\cos\lambda_1 t\sin\theta t + B_1\Sigma\sin\lambda_1 t\sin\theta t \right].$$

If we now let $\lambda_1 = \lambda_2 = \theta = \frac{2\pi m}{N}$, for $m = 1, \ldots, n$ with the help of Appendix Sums (A.2.1)-(A.2.6), EI^2 and EII^2 become

$$\frac{N^4}{4^2}(A_1B_2 - A_2B_1)^2 \quad ,$$

and

$$\frac{N^3\sigma_2^2}{8}\left[A_1^2+B_1^2\right] + \frac{N^3\sigma_1^2}{8}\left[A_2^2+B_2^2\right] - \frac{N^3\rho\sigma_1\sigma_2}{4}\left[A_1A_2+B_1B_2\right],$$

respectively. Hence the expectation of the intensity determinant as given previously in Chapter II is

$$(3.2.5) \qquad E(A_1B_2 - A_2B_1)^2$$

$$= (A_1B_2 - A_2B_1)^2 + \frac{2\sigma_2^2}{N}\left[A_1^2+B_1^2\right] + \frac{2\sigma_2^2}{N}\left[A_2^2+B_2^2\right]$$

$$- \frac{4\sigma_1\sigma_2}{N}\left[A_1A_2 + B_1B_2\right] + \frac{8\sigma_1^2\sigma_2^2}{N^2}(1-\rho^2)$$

In the manner of Chapter I, Section 1.7, it will be instructive to consider the effect of the model period exactly half-way between two trial periods of the integral sub-multiple form N/k. We shall omit all terms in N^{-1} and N^{-2}, considering only the terms I of (3.2.3). Using the Appendix sums (A.5.1)-(A.5.8), I becomes

$$(3.2.5)\quad I^2 \quad \frac{16}{N^4}\left[A_1B_2(-1)(0)-A_1B_2\left(\frac{N}{\pi}\frac{(4k+2)}{4k+1}\right)\left(-\frac{N}{\pi}\frac{4k}{(4k+1)}\right)\right.$$

$$\left.+A_2B_1\left(\frac{N}{\pi}\frac{(4k+2)}{(4k+1)}\right)\left(-\frac{N}{\pi}\frac{(4k)}{(4k+1)}\right)-A_2B_1(-1)(0)\right]^2$$

$$=\left[A_1B_2-A_2B_1\right]^2\frac{N^4}{\pi^4}\left(\frac{4k+2}{4k+1}\right)^2\left(\frac{4k}{4k+1}\right)^2\frac{16}{N^4}$$

$$\approx\frac{16}{\pi^4}\left[A_1B_2-A_2B_1\right]^2\approx 0.16(A_1B_2-A_2B_1)^2 \quad .$$

This is a considerable proportionate reduction in the expectation. It is the square of the proportionate reduction in the univariate case of Section 1.7, and may indicate that the ability of the intensity determinant to detect a non-integral sub-multiple period may be quite low.

3.3. The Expectation of the Intensity Determinant for Different Model Periods and Types of Estimates

It is the purpose of this section to compare the expectations of the intensity determinant under twelve combinations of hypotheses, trial periods, and estimates of the parameters A_1, A_2, B_1, B_2. These twelve situations can be outlined in the following manner:

I. A Real Sinusoidal Component with the Same Period for Both Series.

A. Conventional Estimates (2.3.14)

1. Equal Trial Periods

2. Different Trial Periods

B. Estimates (2.3.17) Based on Different Model Periods

1. Equal Trial Periods

2. Different Trial Periods

II. A Real Sinusoidal Component with a Different Period in Each Series

A. Conventional Estimates (2.3.14)

1. Equal Trial Periods

2. Different Trial Periods

B. Estimates (2.3.17) Based on Different Model Periods

1. Equal Trial Periods

2. Different Trial Periods

III. No Periodic Component in Either Series

A. Conventional Estimates (2.3.14)

1. Equal Trial Periods

2. Different Trial Periods

B. Estimates (2.3.17) Based on Different Model Periods

1. Equal Trial Periods

2. Different Trial Periods

Throughout all expectations involving the estimates (2.3.17) it has been necessary to assume the nuisance parameters $\frac{\rho\sigma_2}{\sigma_1}$ and $\frac{\rho\sigma_1}{\sigma_2}$ as known.

We shall now obtain in detail ten of these expectations. The cases of (I-A-1) and (III-A-1) of the outline have already been given in (3.2.5) and (2.4.14). All of these expectations are evaluated at the point at which the trial period coincides with one or both of the model periods.

(a) Equal model periods, conventional estimates, different periods:

Using (A.2.5) in the basic expression (3.2.2),

$$(3.3.1) \qquad E(\hat{A}_1\hat{B}_2 - \hat{A}_2\hat{B}_1)^2 =$$

$$E\left[4N^{-2}(A_1 \frac{N}{2} + \Sigma\varepsilon_t \cos\theta_1 t)(\Sigma\delta_t \sin\theta_2 t) - 4N^{-2}(\Sigma\delta_t\cos\theta_2 t)(\frac{B_1 N}{2} + \Sigma\varepsilon_t \sin\theta_1 t)\right]$$

$$= 2N^{-1}\sigma_2^2(A_1^2+B_2^2) + N^{-1}8\sigma_1^2\sigma_2^2 \quad .$$

The effect of correlation has been lost, as well as all terms free of the number of observations, N.

(b) Equal model periods, equal trial periods, estimates (2.3.17):

Using the sums (A.2.1.)-(A.2.6), the estimates (2.3.17) become

$$A_1 = 2N^{-1}(\frac{A_1 N}{2} + \Sigma\varepsilon_t\cos\theta t) - 2N^{-1}\frac{\rho\sigma_1}{\sigma_2}(\frac{A_2 N}{2} + \Sigma\delta_t\cos\theta t)$$

$$= (A_1 - \frac{\rho\sigma_1}{\sigma_2}A_2) + 2N^{-1}\Sigma(\varepsilon_t - \frac{\rho\sigma_1}{\sigma_2}\delta_t)\cos\theta t \quad ,$$

$$(3.3.2) \qquad B_1 = 2N^{-1}(\frac{B_1 N}{2} + \Sigma\varepsilon_t\sin\theta t) - 2N^{-1}\frac{\rho\sigma_1}{\sigma_2}(\frac{B_2 N}{2} + \Sigma\delta_t\sin\theta t)$$

$$= (B_1 - \frac{\rho\sigma_1}{\sigma_2}B_2) + 2N^{-1}\Sigma(\varepsilon_t - \frac{\rho\sigma_1}{\sigma_2}\delta_t)\sin\theta t \quad ,$$

$$A_2 = 2N^{-1}\left(\frac{A_2 N}{2} + \Sigma\delta_t \cos\theta t\right) - 2N^{-1}\frac{\rho\sigma_2}{\sigma_1}\left(\frac{A_1 N}{2} + \Sigma\varepsilon_t \cos\theta t\right)$$

$$= \left(A_2 - \frac{\rho\sigma_2}{\sigma_1}A_1\right) + 2N^{-1}\,\Sigma\left(\delta_t - \varepsilon_t\frac{\rho\sigma_2}{\sigma_1}\right)\cos\theta t\ ,$$

$$B_2 = 2N^{-1}\left(\frac{B_2 N}{2} + \Sigma\delta_t \sin\theta t\right) - 2N^{-1}\frac{\rho\sigma_2}{\sigma_1}\left(\frac{B_1 N}{2} + \Sigma\varepsilon_t \sin\theta t\right)$$

$$= \left(B_2 - \frac{\rho\sigma_2}{\sigma_1}B_1\right) + 2N^{-1}\Sigma\left(\delta_t - \frac{\rho\sigma_2}{\sigma_1}\varepsilon_t\right)\sin\theta t\ .$$

Thus,

$$E(\hat{A}_1\hat{B}_2 - \hat{A}_2\hat{B}_1)^2$$

$$\begin{aligned}
= E\Big\{ &\left(A_1 - \frac{\rho\sigma_1}{\sigma_2}A_2\right)\left(B_2 - \frac{\rho\sigma_2}{\sigma_1}B_1\right) - \left(B_1 - \frac{\rho\sigma_1}{\sigma_2}B_2\right)\left(A_2 - \frac{\rho\sigma_2}{\sigma_1}A_1\right) \\
&+ 2N^{-1}\left(A_1 - \frac{\rho\sigma_1}{\sigma_2}A_2\right)\Sigma\left(\delta_t - \frac{\rho\sigma_2}{\sigma_1}\varepsilon_t\right)\sin\theta t \\
&+ 2N^{-1}\left(B_2 - \frac{\rho\sigma_2}{\sigma_1}B_1\right)\Sigma\left(\varepsilon_t - \frac{\rho\sigma_1}{\sigma_2}\delta_t\right)\cos\theta t \\
&- 2N^{-1}\left(B_1 - \frac{\rho\sigma_1}{\sigma_2}B_2\right)\Sigma\left(\delta_t - \varepsilon_t\frac{\rho\sigma_2}{\sigma_1}\right)\cos\theta t \\
&- 2N^{-1}\left(A_2 - \frac{\rho\sigma_2}{\sigma_1}A_1\right)\Sigma\left(\varepsilon_t - \frac{\rho\sigma_1}{\sigma_2}\delta_t\right)\sin\theta t \\
&+ 4N^{-2}\Sigma\left(\varepsilon_t - \frac{\rho\sigma_1}{\sigma_2}\delta_t\right)\cos\theta t\ \Sigma\left(\delta_t - \frac{\rho\sigma_2}{\sigma_1}\varepsilon_t\right)\sin\theta t \\
&- 4N^{-2}\Sigma\left(\varepsilon_t - \frac{\rho\sigma_1}{\sigma_2}\delta_t\right)\sin\theta t\ \Sigma\left(\delta_t - \varepsilon_t\frac{\rho\sigma_2}{\sigma_1}\right)\cos\theta t \Big\}^2
\end{aligned} \tag{3.3.3}$$

$$= \left[(A_1B_2 - A_2B_1)(1-\rho^2) + \rho\frac{(\sigma_2^2 - \sigma_1^2)}{\sigma_2\sigma_1}(A_1B_1 - A_2B_2)\right]^2$$

$$+4\,\sigma_1\sigma_2(1-\rho^2)\left[(A_1-\frac{\rho\sigma_1}{\sigma_2}A_2)(A_2-\frac{\rho\sigma_1}{\sigma_2}A_1)\right.$$

$$\left.+(B_2-\frac{\rho\sigma_2}{\sigma_1}B_1)(B_1-\frac{\rho\sigma_1}{\sigma_2}B_2)\right]$$

$$+8N^{-2}\rho\,\sigma_1^2\sigma_2^2\,(1-\rho^2)^4 \quad .$$

(c) Equal model periods, equal trial periods, Estimates (2.3.17):

In the same manner as the previous expectation,

$$\hat{A}_1 = A_1-\frac{\rho\sigma_1}{\sigma_2}A_2 + 2N^{-1}\,\Sigma(\varepsilon_t-\frac{\rho\sigma_1}{\sigma_2}\delta_t)\cos\Theta_1 t \quad ,$$

$$\hat{B}_1 = (B_1-\frac{\rho\sigma_1}{\sigma_2}B_2) + 2N^{-1}\,\Sigma(\varepsilon_t-\frac{\rho\sigma_1}{\sigma_2}\delta_t)\sin\Theta_1 t \quad ,$$

(3.3.4)

$$\hat{A}_2 = 2N^{-1}\,\Sigma(\delta_t-\frac{\rho\sigma_1}{\sigma_2}\varepsilon_t)\cos\Theta_2 t \quad ,$$

$$\hat{B}_2 = 2N^{-1}\,\Sigma(\delta_t-\frac{\rho\sigma_1}{\sigma_2}\varepsilon_t)\sin\Theta_2 t \quad .$$

$$E\left[\hat{A}_1\hat{B}_2-\hat{A}_2\hat{B}_1\right]^2$$

$$= E\left[2N^{-1}(A_1-\frac{\rho\sigma_1}{\sigma_2}A_2)\,\Sigma(\delta_t-\varepsilon_t\frac{\rho\sigma_2}{\sigma_1})\sin\Theta_2 t\right.$$

$$+4N^{-2}\Sigma(\varepsilon_t-\frac{\rho\sigma_1}{\sigma_2}\delta_t)\cos\Theta_1 t\;\Sigma(\delta_t-\frac{\rho\sigma_2}{\sigma_1}\varepsilon_t)\sin\Theta_2 t$$

(3.3.5)

$$-2N^{-1}(B_1-\frac{\rho\sigma_1}{\sigma_2}B_2)\,\Sigma(\delta_t-\frac{\rho\sigma_2}{\sigma_1}\varepsilon_t)\cos\Theta_2 t$$

$$\left.-4N^{-2}\,\Sigma(\varepsilon_t-\frac{\rho\sigma_1}{\sigma_2}\delta t)\sin\Theta_1 t\;\Sigma(\delta_t-\frac{\rho\sigma_2}{\sigma_1}\varepsilon_t)\cos\Theta_2 t\right]^2$$

$$+4N^{-1}\sigma_2^2(1-\rho^2)\left[(A_1-\rho\frac{\sigma_1}{\sigma_2}A_2)^2+(B_1-\rho\frac{\sigma_1}{\sigma_2}B_2)^2\right]+8N^{-2}\sigma_1^2\sigma_2^2(1-\rho^2)^2 \quad .$$

(d) Different model periods, equal trial periods, Estimates (2.3.14):

Taking the expectation at the point $\theta_1 = \mu \neq \theta_2$,

$$E\left[\hat{A}_1\hat{B}_2-\hat{A}_2\hat{B}_1\right]^2=E\left[4N^{-2}(A_1\Sigma\cos\theta_1 t\cos\mu t+B_1\Sigma\sin\theta_1 t\cos\mu t+\Sigma\varepsilon_t\cos\mu t)\right.$$

$$(A_2\Sigma\cos\theta_2 t\sin\mu t+B_1\Sigma\sin\theta_2 t\sin\mu t+\Sigma\delta_t\sin\mu t)$$

(3.3.6)

$$-4N^{-2}(A_2\Sigma\cos\theta_2 t\cos\mu t+B_2\Sigma\sin\theta_2 t\cos\mu t+\Sigma\delta_t\cos\mu t)$$

$$\left.(A_1\Sigma\cos\theta_1 t\sin\mu t+B_1\Sigma\sin\theta_1 t\sin\mu t+\Sigma\varepsilon_t\sin\mu t)\right]^2$$

$$=E\left[2N^{-1}(A_1-B_1)+4N^{-2}\Sigma\varepsilon_t\cos\mu t\Sigma\delta_t\sin\mu t-4N^{-2}\Sigma\varepsilon_t\sin\mu t\Sigma\delta_t\cos\mu t\right]^2$$

$$=4N^{-2}(A_1-B_1)^2+8N^{-2}\sigma_1^2\sigma_2^2-32N^{-2}E\left[\Sigma\varepsilon_t\cos\mu t\Sigma\delta_t\sin\mu t\Sigma\varepsilon_t\sin\mu t\Sigma\delta_t\cos\mu t\right]$$

$$=4N^{-2}(A_1-B_1)^2+8N^{-2}\sigma_1^2\sigma_2^2(1-\rho^2) \ .$$

(e) Different model periods, different trial periods, Estimates (2.3.14):

Evaluating the expectation at the point $\theta_1 = \mu_1$, $\theta_2 = \mu_2$,

$$E\left[\hat{A}_1\hat{B}_2-\hat{A}_2\hat{B}_1\right]^2$$

$$=E\left[4N^{-2}(A_1\Sigma\cos\theta_1 t\cos\mu_1 t+B_1\Sigma\sin\theta_1 t\cos\mu_1 t+\Sigma\varepsilon_t\cos\mu_1 t)\right.$$

$$\cdot(A_2\Sigma\cos\theta t\sin\mu_2 t + B_2\Sigma\sin\theta_2 t\sin\mu_2 t + \Sigma\delta_t\sin\mu_2 t)$$

(3.3.7)

$$-4N^{-2}(A_2\Sigma\cos\theta_2 t\cos\mu_2 t + B_2\Sigma\sin\theta_2 t\cos\mu_2 t + \Sigma\delta_t\cos\mu_2 t)$$

$$\cdot(A_1\Sigma\cos\theta_1 t\sin\mu_1 t + B_1\Sigma\sin\theta_1 t\sin\mu_1 t + \Sigma\varepsilon_t\sin\mu_1 t)\,]^2$$

$$=E[(A_1B_2-A_2B_1)+2N^{-1}(A_1\Sigma\delta_t\sin\mu_2 t+ B_2\Sigma\varepsilon_t\cos\mu_1 t$$

$$- A_2\Sigma\varepsilon_t\sin\mu_1 t- B_1\Sigma\delta_t\cos\mu_2 t)$$

$$+4N^{-2}(\Sigma\varepsilon_t\cos\mu_1 t\,\Sigma\delta_t\sin\mu_2 t-\Sigma\delta_t\cos\mu_2 t\;\Sigma\varepsilon_t\sin\mu_1 t)\quad]^2$$

$$= (A_1B_2-A_2B_1)^2 + 8N^{-2}\sigma_1^2\sigma_2^2 \quad .$$

The effect of correlation has been lost in this choice of trial and model periods.

(f) Different model periods, equal trial periods, Estimates (2.3.17):

Expecting at the point $\theta\,\mu_1 \neq \mu_2$ with estimates given by

$$\hat{A}_1 = A_1 + 2N^{-1}\,\Sigma(\varepsilon_t - \frac{\rho\sigma_1}{\sigma_2}\,\delta_t)\cos\theta t \quad ,$$

$$\hat{B}_1 = B_1 + 2N^{-1}\,\Sigma(\varepsilon_t - \frac{\rho\sigma_1}{\sigma_2}\,\delta_t)\sin\theta t \quad .$$

(3.3.8)

$$\hat{A}_2 = -\frac{\rho\sigma_2}{\sigma_1}A_1 + 2N^{-1}\,\Sigma(\delta_t - \frac{\rho\sigma_2}{\sigma_1}\,\varepsilon_t)\cos\theta t \; ,$$

$$\hat{B}_2 = -\frac{\rho\sigma_2}{\sigma_1} + 2N^{-1}\,\Sigma(\delta_t - \frac{\rho\sigma_2}{\sigma_1}\,\varepsilon_t)\sin\theta t \; ,$$

(3.3.9) $$E[\hat{A}_1\hat{B}_2-\hat{A}_2\hat{B}_1]^2=4A_1^2B_1^2\,\rho\,\frac{\sigma_2^2}{\sigma_1^2} + 2N^{-1}\sigma_2^2(A_1^2+B_1^2)[(1-\rho^4)+2(1-\rho^2)]$$

$$+ 8N^{-2}\sigma_1^2\sigma_2^2(1-\rho^2)^3 \quad .$$

(g) Different model periods, different trial periods, Estimates (2.3.17):

Let the model Periods be μ_1 μ_2. We wish the expectation at the point $\theta_1 = \mu_1$, $\theta_2 = \mu_2$. Using the necessary Appendix sums, the estimates A_1, A_2, B_1, B_2 become

(3.3.10)
$$\hat{A}_1 = \frac{2}{N}\left[\frac{A_1 N}{2} + \Sigma\varepsilon_t \cos\theta_1 t\right] - \frac{2}{N}\,\frac{\rho\sigma_1}{\sigma_2}(\Sigma\delta_t\cos\theta_1 t) \quad ,$$
$$\hat{B}_1 = B_1 + \frac{2}{N}\left[\Sigma(\varepsilon_t - \delta_t\frac{\rho\sigma_1}{\sigma_2})\sin\theta_1 t\right] \quad ,$$
$$\hat{A}_2 = A_2 + \frac{2}{N}\left[\Sigma(\delta_t - \varepsilon_t\frac{\rho\sigma_2}{\sigma})\cos\theta_2 t\right] \quad ,$$
$$\hat{B}_2 = B_2 + \frac{2}{N}\left[\Sigma(\delta_t - \varepsilon_t\frac{\rho\sigma_2}{\sigma_1})\sin\theta_2 t\right] \quad .$$

Inserting these in the intensity determinant and expecting, after some simplification we have

(3.3.11)
$$E\left[\hat{A}_1\hat{B}_2 - \hat{A}_2\hat{B}_1\right]^2 = (A_1B_2 - A_2B_1)^2$$
$$+ \frac{4}{N^2}\left[\frac{A_1^2N\sigma_2^2}{2}(1-\rho^2) + \frac{B_2^2N\sigma_1^2}{2}(1-\rho^2)\right.$$
$$\left. + \frac{N\sigma_1^2A_2^2}{2}(1-\rho^2) + \frac{N}{2}\sigma_2^2B_1^2(1-\rho^2)\right]$$
$$+ \frac{16}{N^4}\left[\frac{N^2}{4}\sigma_1^2\sigma_2^2(1-\rho^2)^2 + \frac{N^2}{4}\sigma_1^2\sigma_2^2(1-\rho^2)\right]$$

$$=(A_1B_2-A_2B_1)^2+\frac{2}{N}(1-\rho^2)\left[\sigma_2^2(A_1^2+B_1^2)+\sigma_1^2(A_2^2+B_2^2)\right]$$

$$+\frac{8\sigma_1^2\sigma_2^2}{N}(1-\rho^2)^2 \quad .$$

(h) Null hypothesis of no periodic component in either series, different trial periods, Estimates (2.3.14):

The estimates become simply

(3.3.12) $$\hat{A}_1=2N^{-1}\Sigma\varepsilon_t\cos\Theta_1 t \quad , \quad \hat{A}_2=2N^{-1}\Sigma\delta_t\cos\Theta_2 t \quad ,$$

$$\hat{B}_1=2N^{-1}\Sigma\varepsilon_t\sin\Theta_1 t \quad , \quad \hat{B}_2=2N^{-1}\Sigma\delta_t\sin\Theta_2 t \quad .$$

Hence,

(3.3.13) $$E\left[\hat{A}_1\hat{B}_2-\hat{A}_2\hat{B}_1\right]^2=8N^{-2}\sigma_1^2\sigma_2^2 \quad .$$

The use of different trial periods and conventional estimates under the null hypothesis has removed the effect of inter-series correlation from this expectation.

(i) Null hypothesis of no periodic component in either series, equal trial periods, Estimates (2.3.17):

For Θ the common trial period, the expected value of the intensity determinant under the above conditions is

(3.3.14) $E\left[\hat{A}_1\hat{B}_2-\hat{A}_2\hat{B}_1\right]^2$

$$=E\left[4N^{-2}\left[\Sigma(\varepsilon_t-\frac{\rho\sigma_1}{\sigma_2}\delta_t)\cos\Theta t\right]\left[\Sigma(\delta_t-\frac{\rho\sigma_2}{\sigma_1}\varepsilon_t)\sin\Theta t\right]\right.$$

$$-4N^{-2}\left[\Sigma(\delta_t-\frac{\rho\sigma_2}{\sigma_1}\varepsilon_t)\cos\theta t\right]\left[\Sigma(\varepsilon_t-\frac{\rho\sigma_1}{\sigma_2}\delta_t)\sin\theta t\right]^2$$

$$= 8N^{-2}\sigma_1^2\sigma_2^2(1-\rho^2)^3 \quad .$$

The usual expectation under the null hypothesis is decreased by the factor $(1-\rho^2)^2$.

(j) Null hypothesis of no periodic component in either series, different trial periods, Estimates (2.3.17):

Write the estimates (2.3.25)-(2.3.28) as

(3.3.15)

$$\hat{A}_1 = \frac{2}{N}\left[\Sigma(u_t-\frac{\rho\sigma_1}{\sigma_2}v_t)\cos\lambda_1 t\right]=\frac{2}{N}\Sigma\varepsilon_t\cos\lambda_1 t = x_1 \quad ,$$

$$\hat{B}_1 = \frac{2}{N}\left[\Sigma(u_t-\frac{\rho\sigma_1}{\sigma_2}v_t)\sin\lambda_1 t\right]=\frac{2}{N}\Sigma\varepsilon_t\sin\lambda_1 t = x_2 \quad ,$$

$$\hat{A}_2 = \frac{2}{N}\left[\Sigma(v_t-\frac{\rho\sigma_2}{\sigma_1}u_t)\cos\lambda_2 t\right]=\frac{2}{N}\Sigma\delta_t\cos\lambda_2 t = x_3 \quad ,$$

$$\hat{B}_2 = \frac{2}{N}\left[\Sigma(v_t-\frac{\rho\sigma_2}{\sigma_1}u_t)\sin\lambda_2 t\right]=\frac{2}{N}\Sigma\delta_t\sin\lambda_2 t = x_4 \quad ,$$

Inserting these into the determinant $(\hat{A}_1\hat{B}_2-\hat{A}_2\hat{B}_1)^2$, and expecting, we wish the values of $Ex_1^2x_2^2$, $Ex_3^2x_4^2$, $Ex_1x_2x_3x_4$. Since the following relations:

(3.3.16)

$$Ex_1x_2 = 0, \qquad Ex_2x_3 = 0 \quad ,$$

$$Ex_1x_3 = 0, \qquad Ex_2x_4 = 0 \quad ,$$

$$Ex_1x_4 = 0, \qquad Ex_3x_4 = 0 \quad ,$$

hold from the orthogonality of the trigonometric sums involved we have:

(3.3.17) $$Ex_1^2x_2^2 = Ex_3^2x_4^2 = \frac{4}{N^2}\sigma_1^2\sigma_2^2(1-\rho^2) \quad ,$$

$$Ex_1x_2x_3x_4 = 0 \quad .$$

Thus, under the conditions imposed above,

(3.3.18) $$E(A_1B_2 - A_2B_1)^2 = \frac{8}{N^2}\sigma_1^2\sigma_2^2(1-\rho^2)^2 \quad .$$

The usual null hypothesis expectation has been decreased by the factor $(1-\rho^2)$.

APPENDIX

CERTAIN TRIGONOMETRIC SUMS

1.1. Introduction

Throughout this thesis it has been necessary to evaluate many trigonometric sums of the type

$$\sum_{t=1}^{N} \rho^{t} \cos^{r} \frac{2\pi m}{N} t \sin^{s} \frac{2\pi n}{N} t \quad ,$$

where r or s is usually small, if not zero, and ρ is usually 1. These results have long been known to workers in applied mathematics, but a detailed table for certain values of r, s, m, and n, is given in this Appendix because of their frequent use in Chapters II and III.

The sum

$$\sum_{t=a}^{b} e^{i\theta t} \rho^{t} \tag{A.1.1}$$

is simply a geometric progression, and may be written as

$$\frac{\rho^{b+a} e^{i\theta(b+a)} - \rho^{a} e^{i\theta a}}{\rho e^{i\theta} - 1} \quad . \tag{A.1.2}$$

Multiplying the numerator and denominator by the complex conjugate of the latter, this becomes

$$\frac{\rho^{b+1+a} e^{i\theta(b+a-1)} - \rho^{b+a} e^{i\theta(b+a)} - \rho^{a+1} e^{i\theta(a-1)} + \rho^{a} e^{i\theta a}}{\rho^{2} - \rho e^{i\theta} - \rho e^{-i\theta} + 1} \tag{A.1.3}$$

Taking real and imaginary parts separately, we have

(A.1.4)

$$\sum_{t=a}^{b} \rho^t \cos\theta t = \frac{\rho^{b+1+a}\cos(b+a-1)\theta - \rho^{b+a}\cos(b+a)\theta - \rho^{a+1}\cos(a-1)\theta + \rho^a \cos\theta a}{\rho^2 - 2\rho\cos\theta + 1} .$$

(A.1.5)

$$\sum_{t=a}^{b} \rho^t \sin\theta t = \frac{\rho^{b+1+a}\sin(b+a-1)\theta - \rho^{b+a}\sin(b+a) - \rho^{a+1}\sin(a-1)\theta + \rho^a \sin\theta a}{\rho^2 - 2\rho\cos\theta + 1} .$$

These general expressions form the basis for the following table of summations.

A.2. A Table of Trigonometric Sums

(a) $\theta = \frac{2\pi k}{N}$, $a = 1$, $b = N$, $k = 1,2,\dots N-1$;

(A.2.1) $$\sum_{t=1}^{N} \cos\theta t = 0$$

(A.2.2) $$\sum_{t=1}^{N} \sin\theta t = 0$$

(A.2.3) $$\sum_{t=1}^{N} \sin t\theta \cos t\theta = 0$$

(A.2.4) $$\sum_{t=1}^{N} \sin^2 t\theta = \frac{N}{2}$$

(A.2.5) $$\sum_{t=1}^{N} \cos^2 t\theta = \frac{N}{2}$$

(A.2.6) $$\sum_{t=1}^{N} \sin\frac{2\pi m}{N}t \sin\frac{2\pi n}{N} t = 0,\ m \neq n$$

(A.2.7) $\sum_{t=1}^{N} \cos \frac{2\pi m}{N}t \cos \frac{2\pi n}{N}t = 0, \; m \neq n$

(b) $\theta = \frac{2\pi k}{N}$, $a = 2$, $b = N$, $\rho = 1$; $k=1,2,\ldots, N-1$:

(A.3.1) $\sum_{t=2}^{N} \sin(t-1)\theta \sin t\theta = \frac{N}{2} \cos \theta$

(A.3.2) $\sum_{t=2}^{N} \sin(t-1)\theta \cos t\theta = \frac{-N}{2} \sin \theta$

(A.3.3) $\sum_{t=2}^{N} \cos^2(t-1)\theta = \frac{N-2}{2}$

(A.3.4) $\sum_{t=2}^{N} \cos(t-1)\theta \cos t\theta = \frac{(N-2)}{2} \cos \theta$

(A.3.5) $\sum_{t=2}^{N} \sin^2(t-1)\theta = \frac{N}{2}$

(A.3.6) $\sum_{t=2}^{N} \cos(t-1)\theta \sin(t-1)\theta = 0$

(A.3.7) $\sum_{t=2}^{N} \cos(t-1)\theta \sin t\theta = \frac{(N-2)}{2} \sin \theta$

(A.3.8) $\sum_{t=2}^{N} \sin(t-1)\theta = 0$

(c) $\theta = \frac{2nk}{N}$, $b = N$, ρ a real number, $k = 1,2,\ldots, N-1$:

(A.4.1) $\sum_{t=2}^{N} \rho^t \cos t\theta = \frac{\rho^{N+a+1} \cos(a-1)\theta - \rho^{N+a} \cos a\theta}{\rho^2 - 2\rho\cos\theta + 1}$

$$\frac{-\rho^{a+1}\cos(a-1)\theta+\rho^{a}\cos a\theta}{\rho^{2}-2\rho\cos\theta+1}$$

For a = 1,

(A.4.2) $$\sum_{t=1}^{N}\rho^{t}\cos t\theta=\frac{\rho^{N+2}-\rho^{2}-\rho^{N+1}\cos\theta+\rho\cos\theta}{\rho^{2}-2\rho\cos\theta+1},$$

(A.4.3) $$\sum_{t=2}^{N}\rho^{t}\sin t\theta=\frac{\rho^{N+1+a}\sin(a-1)\theta-\rho^{N+2}\sin a\theta}{\rho^{2}-2\rho\cos\theta+1}$$

$$\frac{-\rho^{a+1}\sin(a-1)\theta+\rho^{a}\sin a\theta}{\rho^{2}-2\rho\cos\theta+1}.$$

For a = 1,

(A.4.4) $$\sum_{t=1}^{N}\rho^{t}\sin t\theta=\frac{-\rho^{N+1}\sin\theta+\rho\sin\theta}{\rho^{2}-2\rho\cos\theta+1}.$$

(d) $\theta=\frac{\pi k}{N}$, $\lambda=\frac{2\pi}{N}(k+\frac{1}{2})$, $a=1$, $b=N$, $\rho=1$, $k=1, 2, \ldots, N-1$:

(A.5.1) $$\sum_{t=1}^{N}\sin t\theta=\frac{\sin\theta}{1-\cos\theta}, \quad k \text{ odd}.$$

$$=0, \quad k \text{ even}.$$

(A.5.2) $$\sum_{t=1}^{N}\cos t\theta=-1, \quad k \text{ odd}.$$

$$=0, \quad k \text{ even}.$$

(A.5.3) $$\sum_{t=1}^{N}\cos t=-1.$$

(A.5.4) $\sum_{t=1}^{N} \cos t(\lambda+2\theta) = -1$.

(A.5.5) $\sum_{t=1}^{N} \cos t\lambda \cos 2t\theta = -1$.

(A.5.6) $\sum_{t=2}^{N} \sin t\lambda \sin 2t\theta = 0$.

(A.5.7) $$\sum_{t=1}^{N} \sin t\lambda \cos 2t\theta = \frac{1}{2}\left[\frac{\sin \frac{\pi(4k+1)}{N}}{1-\cos \frac{\pi(4k+1)}{N}} + \frac{\sin \frac{\pi}{N}}{1-\cos \frac{\pi}{N}} \right]$$

$$\approx \frac{N}{\pi}\left[\frac{4k+2}{4k+1} \right] .$$

(A.5.8) $$\sum_{t=1}^{N} \sin 2t\theta \cos t\lambda = \frac{1}{2}\left[\frac{\sin \frac{\pi(4k+1)}{N}}{1-\cos \frac{\pi(4k+1)}{N}} - \frac{\sin \frac{\pi}{N}}{1-\cos \frac{\pi}{N}} \right]$$

$$\approx \frac{-N4k}{\pi 4k+1} .$$

(e) $\theta = \frac{2\pi k}{N}$, $\lambda = \frac{\pi(4k+1)}{2N}$, $a=1$, $b=n$, $\rho=1$, $k = 0, 1, \ldots, N-1$:

(A.6.1) $$\sum_{t=1}^{N} \sin t\lambda = \frac{1}{2} + \frac{\sin \lambda}{2(1-\cos\lambda)} \approx \frac{1}{2} + \frac{2N}{(4k+1)} .$$

(A.6.2) $$\sum_{t=1}^{N} \cos t\lambda = \frac{-1}{2} + \frac{\sin \lambda}{2(1-\cos \lambda)} \approx \frac{1}{2} + \frac{2N}{(4k+1)} .$$

(A.6.3) $$\sum_{t=1}^{N} \sin t\lambda \cos t\theta = \frac{1}{2}\left[1 + \frac{\sin \frac{\pi}{2N}}{2(1-\cos \frac{\pi}{2N})} + \frac{\sin \frac{\pi(8k+1)}{2N}}{2(1-\cos \frac{\pi(8k+1)}{2N})} \right]$$

$$\approx \quad \frac{1}{2} + \frac{N}{\pi}\left[\frac{8k+2}{8k+1}\right] \approx \frac{N}{\pi} \quad .$$

$$(A.6.4) \quad \sum_{t=1}^{N} \cos t\lambda \cos t\theta = \frac{1}{2}\left[-1+\frac{\sin\frac{\pi}{2N}}{2(1-\cos\frac{\pi}{2N})}+\frac{\sin\frac{\pi(\theta k+1)}{2N}}{2(1-\cos\frac{\pi(8k+1)}{2N})}\right]$$

$$\approx \quad -\frac{1}{2} + \frac{N}{\pi}\left[\frac{8k+2}{8k+1}\right] \approx \frac{N}{\pi} \quad .$$

$$(A.6.5) \quad \sum_{t=1}^{N} \sin t\lambda \sin t\theta = \frac{1}{2}\left[-\frac{1}{2}+\frac{\sin\frac{\pi(8k+1)}{2N}}{2(1-\cos\frac{\pi(8k+1)}{2N})}+\frac{1}{2}-\frac{\sin\frac{\pi}{2N}}{2(1-\cos\frac{\pi}{2N})}\right]$$

$$\approx \quad \frac{N}{\pi}\,\frac{8k}{8k+1} \approx \frac{N}{\pi} \quad .$$

$$(A.6.6) \quad \sum_{t=1}^{N} \cos t\lambda \sin t\theta = \frac{1}{2}\left[\frac{1}{2}+\frac{1 \sin\frac{\pi(8k+1)}{2N}}{2(1-\cos\frac{\pi(8k+1)}{2N})}-\frac{1}{2}-\frac{\sin\frac{\pi}{2N}}{2(1-\cos\frac{\pi}{2N})}\right]$$

$$\approx \quad -\frac{N}{\pi}\,\frac{8k}{(8k+1)} \approx -\frac{N}{\pi} \quad .$$

It is interesting to note that the last four sums are all approximately equal to $N\pi^{-1}$ in absolute value. The range of values of k has been chosen to only supply the needed sums for Section 1.7.

BIBLIOGRAPHY

[1] Bartlett, M. S., An Introduction to Stochastic Processes, Cambridge, Cambridge University Press, 1955.

[2] Cochran, William G., "The Distribution of the Largest of a Set of Estimated Variances as a Fraction of Their Total," Annals of Eugenics, Vol XI(1941), 47-52.

[3] Davis, H. T., The Analysis of Economic Time Series, Bloomington, Principia Press, 1941.

[4] Fisher, Ronald A., "The General Sampling Distribution of the Multiple Correlation Coefficient." Proceedings of the Royal Society of London, Vol. 121 (1928), 654.

[5] Fisher, Ronald A., "Tests of Significance in Harmonic Analysis," Proceedings of the Royal Society of London, Vol. 125 (1929), 54.

[6] Fisher, R. A., "The Sampling Distribution of Some Statistics Obtained from Non-linear Equations," Annals of Eugenics, Vol. IX (1939), 238-249.

[7] Fisher, R. A., "On the Similarity of the Distributions found for the Test of Significance in Harmonic Analysis, and in Stevens' Problem in Geometrical Probability," Annals of Eugenics, Vol. X (1940), 14-17.

[8] Hartley, H. O., "Tests of Significance in Harmonic Analysis," Biometrika, Vol. XXXVI (1949), 194-201.

[9] Hotelling, Harold, "A Generalized T test and Measure of Multivariate Dispersion," Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Vol. II (1951), 23-41.

[10] Nanda, D. N., "Distribution of a Root of a Determinantal Equation," Annals of Mathematical Statistics, Vol. XIX (1948), 340-350.

[11] Pillai, K.C.S., On Some Distribution Problems in Multivariate Analysis (Unpublished Thesis), Chapel Hill, N. C., 1954.

[12] Rao, C. R., Advanced Statistical Methods in Biometric New York, Wiley, 1952.

[13] Schuster, Sir Arthur, "On the Investigation of Hidden Periodicities with Application to a Supposed 26 Day Period of Metereological Phenomena," Terrestrial Magnetism, Vol. III (1898), 13.

[14] Starkey, Daisy M., "The Distribution of the Multiple Correlation Coefficient in Periodogram Analysis," Annals of Mathematical Statistics, Vol. X (1939), 327.

[15] Stevens, W. L., "Solution to a Geometrical Problem in Probability," Annals of Eugenics, Vol. IX (1939), 315-320.

[16] Stumpff, K., "Ermittlung und Realitat von Periodizitaten," Korrelationsrechnung, Berlin, Borntraeger, 1940.

[17] Walker, Gilbert, "On the Criterion for the Reality of Relationships or Periodicities," Indian Meterological Memoirs, Vol. XXI, Part 9.

[18] Whittle, Peter, Hypothesis Testing in Time Series Analysis. Uppsala, Almqvist and Wicksell, 1951.

[19] Wilks, S. S., "Certain Generalizations in the Analysis of Variance," Biometrika, XXIV (1932), 471-494.

[20] Wold, Herman, A Study in the Analysis of Stationary Time Series. Stockholm, Almqvist and Wicksell, 1954.